UNIVERSIDAD AUTÓNOMA DE MADRID

ESCUELA POLITÉCNICA SUPERIOR

DYNAMIC SIGNATURE VERIFICATION FOR PORTABLE DEVICES

–TRABAJO DE FIN DE MÁSTER–

Author: Marcos Martínez Díaz
Ingeniero de Telecomunicación, UAM

A thesis submitted for the degree of

Máster en Ingeniería Informática y de Telecomunicación

Madrid, November 26, 2008

Escuela Politécnica Superior
Universidad Autónoma de Madrid (UAM), Spain

Dynamic Signature Verification
for Portable Devices

Autor: **Marcos Martínez Díaz**
Ingeniero de Telecomunicación (UAM)

Director: **Julián Fiérrez Aguilar**
Doctor Ingeniero de Telecomunicación (UPM)
Universidad Autónoma de Madrid

Tutor: **Javier Ortega García**
Doctor Ingeniero de Telecomunicación (UPM)
Universidad Autónoma de Madrid

Fecha: 26 de noviembre de 2008

Tribunal: **Javier Ortega García**
Universidad Autónoma de Madrid

Joaquín González Rodríguez
Universidad Autónoma de Madrid

Juan A. Sigüenza Pizarro
Universidad Autónoma de Madrid

Calificación:

The research described in this work was carried out in the Biometric Recognition Group–ATVS at the Dept. of Ingeniera Informática, Escuela Politécnica Superior, Universidad Autónoma de Madrid. This work has been supported by the Spanish Ministry of Education under project TEC2006-13141-C03-03 and by the European Network of Excellence Biosecure (IST-2002-507634).

Acknowledgements

I would like to thank my advisors Prof. Julian Fierrez and Prof. Javier Ortega-Garcia for their guidance in the last years. I also wish to thank Manuel Freire, Javier Galbally, Fernando Alonso-Fernandez, Pedro Tome and Daniel Ramos-Castro just to mention some of the colleagues from the the ATVS - Biometric Recognition Group that have helped and supported me.

Abstract

The proliferation of handheld devices such as PDAs and smartphones represents a new scenario for automatic signature verification. Traditionally, research on signature verification has been carried out employing signatures acquired using digitizing tablets or Tablet-PCs. In this work we study the effects of the mobile acquisition conditions and we analyze the considerations that must be taken in the new handheld scenario. The analysis is performed with two state-of-the art systems: one feature-based and another function-based. Two perspectives are taken on the analysis. First, a comparison of the handheld and the pen tablet scenario from a statistical point of view based on class separability measures is performed. Second, feature selection is carried out to investigate the most relevant feature combinations in both systems and both scenarios. Results confirm that the lack of pen-up trajectory information on the handheld scenario negatively affects the verification performance, especially against skilled forgeries. Finally, a system based on fusion of the global and the local system is presented, with promising results.[1]

[1]Part of this work will be published in the Proceedings of the 19th International Conference in Pattern Recognition, ICPR 2008. The accepted paper is included in Appendix A.

Contents

List of Figures

Chapter 1

Introduction

The proliferation of touchscreen-enabled devices represents many new promising scenarios and applications for signature verification. Automatic signature verification is a challenging task *per se*, as it must face a notable variability among signatures from the same individual and the risk of highly skilled forgers which, due to their unpredictable nature, are not completely possible to model during the design of a verification system. This work is focused on automatic person authentication using signature as a biometric trait. Dynamic signature verification for portable devices is studied and an analysis of its specificities compared to traditional signature verification systems based on digitizing tablets is performed.

Within biometrics, signature is one of the most socially accepted biometric traits, as it has been used in financial and legal transactions for centuries (Fierrez and Ortega-Garcia, 2007b; Plamondon and Lorette, 1989). In the current era of electronic services and ubiquitous access to information, secure access control and user authentication are common tasks which are usually performed with tokens or passwords. In this field, biometrics has become a focus of interest as it uses anatomical (e.g. fingerprint, iris) or behavioral (e.g. voice, signature) traits to authenticate a user (Jain *et al.*, 2004). These traits cannot be easily stolen or forgotten. It is now common to observe fingerprint verification systems in portable electronic devices (e.g. handhelds), face recognition systems for border control purposes and iris verification in some airports (e.g. United Arab Emirates).

Biometric authentication has gathered an increasing research and commercial interest in the last few years (Jain *et al.*, 2006) as it represents a convenient and secure means of person authentication.

1.1 Biometrics

Biometrics are generally used for *identification* or *verification* purposes (Jain *et al.*, 2004). In the former mode of operation, the biometric trait that individuals present to the system is used to determine which one of the enrolled users in the database they are, leading to a $1 : N$ comparison, where N is the number of users in the database. In the latter, the biometric trait is used to authenticate an individual claiming to be a specific user, which is performed by a $1 : 1$ comparison between the provided biometric trait and the enrolled data of the claimed user. Throughout this work, we will address the problem of verification, also known as authentication.

Verification systems are essentially two-class classifiers, which produce an *accept* or *reject* decision when a biometric trait along with a user identity are presented to the system. Usually, verification is based on a decision threshold. If the similarity (or match score) between the provided trait and the model from the claimed user is higher than a specific threshold, the user is accepted by the system. On the contrary, the user is rejected. In this context, verification systems face two type of errors: False Acceptance (FA) and False Rejection (FR). False Acceptance is produced when a user that falsely claims to be another user is accepted by the system as being the genuine user. False Rejection means that a genuine user is rejected by the system as being an impostor. Given a population of genuine users and impostors and a series of verification trials, the False Acceptance Rate (FAR) and False Rejection Rate (FRR) of the biometric verification system at hand can be computed for any decision threshold.

A common measure to compare the performance of biometric systems is the *Equal Error Rate* (EER). It is computed as the system error rate when the decision threshold is set to satisfy that $FAR = FRR$. Due to the fact that the output of a verification system is in general a binary decision (i.e. accept/reject), the performance of a biometric system is usually represented by a Receiver Operating Characteristic (ROC) or a Detection Error Trade-off (DET) plot (Martin *et al.*, 1997). These plots allow an easy comparison between different systems at any decision threshold.

1.1.1 Biometric Modalities

Several biometric modalities have been proposed in the last decades (Jain *et al.*, 2004). These can be based on physical and behavioral traits depending on their nature. Physical traits are related to anatomical properties of an individual, and include fingerprint, face, iris and hand geometry among others. Behavioral traits refer to how an individual performs an action, and include voice, signature and gait among the most typical. Some examples of popular biometric traits are presented in Fig. 1.1.

Biometric modalities can be further classified by other measures such as the following:

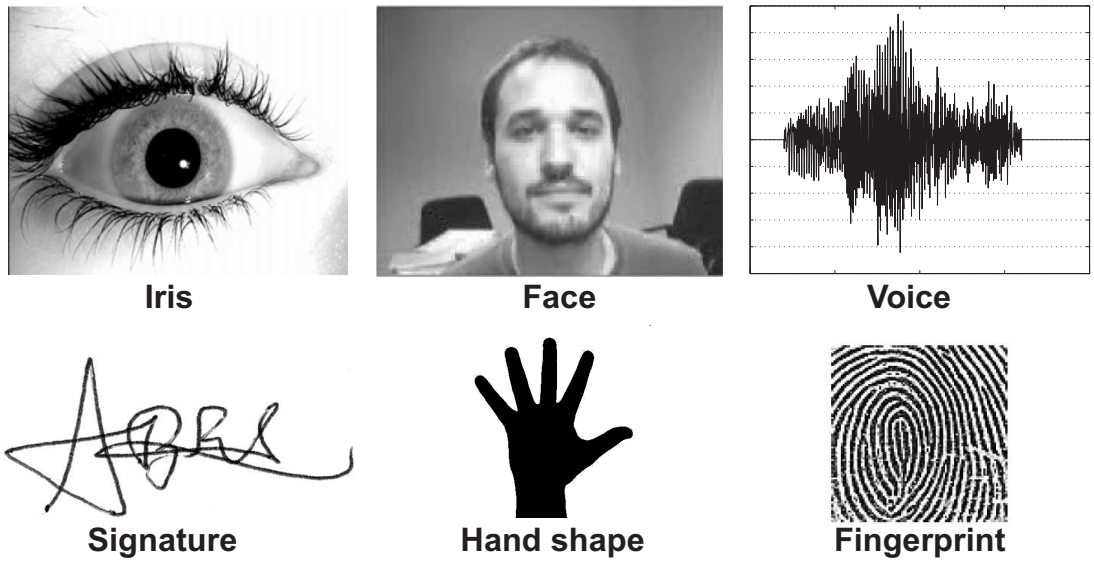


Figure 1.1: *Examples of biometric traits.*

- *Universality*, which states if every person has this biometric.
- *Distinctiveness*, related to the discriminative power between different individuals of a biometric modality.
- *Permanence*, which is higher if the traits are invariant along periods of time.
- *Collectability*, which refers to how easy is to acquire the biometric trait.
- *Performance*, related to the speed, or accuracy of systems based on a given biometric.
- *Acceptability*, related to the social perception of the biometric modality.
- *Circumvention*, which refers to the resilience against attacks to security systems based on the biometric.

A comparison between some popular biometrics based on the aforementioned measures is presented in Table 1.1. As can be seen, no specific biometric outperforms the rest of them on every category. Consequently, the choice of a modality will depend on the application it is intended to be used for.

1.2 Signature Verification

Signatures have been used since centuries to validate documents and transactions. Therefore, signature is one of the most socially accepted among all biometric traits. In the last few decades,

Table 1.1: *Qualitative comparison of popular biometric modalities. H, M and L denote High, Medium, and Low respectively. Adapted from (Jain et al., 2004).*

Biometric	Universality	Distinctiveness	Permanence	Collectability	Performance	Acceptability	Circumvention
Face	H	L	M	H	L	H	H
Fingerprint	M	H	H	M	H	M	M
Gait	M	L	L	H	L	H	M
Hand Geometry	M	M	M	H	M	M	M
Iris	H	H	H	M	H	L	L
Keystroke	L	L	L	M	L	M	M
Signature	L	L	L	H	L	H	H
Voice	M	L	L	M	L	H	H

digitizing devices have allowed to perform machine-based signature verification, which has been an intense research field among the biometric and handwriting recognition research communities.

Despite its wide acceptance, automatic signature verification is still a challenging task. This can be corroborated by the variety of research works conducted in the last years (Fierrez and Ortega-Garcia, 2007b; Leclerc and Plamondon, 1994; Plamondon and Lorette, 1989; Plamondon and Srihari, 2000). One of the main challenges in signature verification is related to the signature variability. While signatures from the same user show considerable differences between different captures (high *intra-class* variability), skilled forgers can perform signatures with high resemblance to the user's signature (low *inter-class* variability). Moreover, when a system is designed, only a fraction of information about skilled forgeries can be obtained as forgers with unexpected skills can appear at any time once the system has been deployed.

Two main classes of signature verification systems exist depending on the information extracted from the signature. Off-line systems use only the signature image, while on-line or dynamic systems employ digitized time functions of the signature.

Off-line signature verification systems use static signature images, which may have been scanned or acquired using a camera, to perform verification. The approaches taken for off-line signature verification have been heterogeneous. Some authors focus on global features using image or shape-oriented pattern recognition techniques (Sabourin, 1997) while others use local features, relying on stroke, texture and structural information (Ammar *et al.*, 1990; Guo *et al.*,

1997). Some approaches combine both global and local features (Fierrez-Aguilar *et al.*, 2004; Huang and Yan, 1997).

On-line or **dynamic** systems use captured signature time-functions. These functions are obtained using digitizer tablets or touchscreens (e.g. Tablet-PCs, smart phones, etc.). Traditionally, dynamic systems have presented a better performance than off-line systems as more levels of information than the signature static image are available (Plamondon and Lorette, 1989). This is the approach considered in this work, and will be described in the following chapters.

1.3 Applications of Signature Verification on Handheld Devices

Touchscreen mobile devices such as smart phones or PDAs provide an appropriate computing platform for signature verification. In fact, commercial devices already provide handwritten character recognition as a text input alternative (Anquetil and Bouchereau, 2002; Ballagas *et al.*, 2006).

Signature verification can be used for a wide range of applications. Among them, we cite the following:

Payments in commercial environments: the signature is used to validate a payment that is performed via WiFi, UMTS, GPRS, or other mobile network. This enables ubiquitous access to commercial transactions. Currently, signatures are not always visually verified by the cashier, so automatic verification could provide higher security levels.

Legal transactions: legal documents or certificates are signed by the user adding additional security as the signature is verified. This can be a convenient user validation scheme for e-government applications. Using on-line signature verification, the protection against repudiation of signed documents is even increased over traditional signature.

User login: the signature is used to login into a local or remote system as an access control measure (e.g. bank account, personal records, etc.), instead of traditional methods such as PINs or passwords.

Customer validation: a customer is validated by its signature. A client that receives a service or a delivery (e.g. a parcel) signs in a mobile device carried by the deliverer or service provider to certify his conformity.

Cryptobiometrics: signature is used as a cryptographic key (Freire-Santos *et al.*, 2006) that identifies the user.

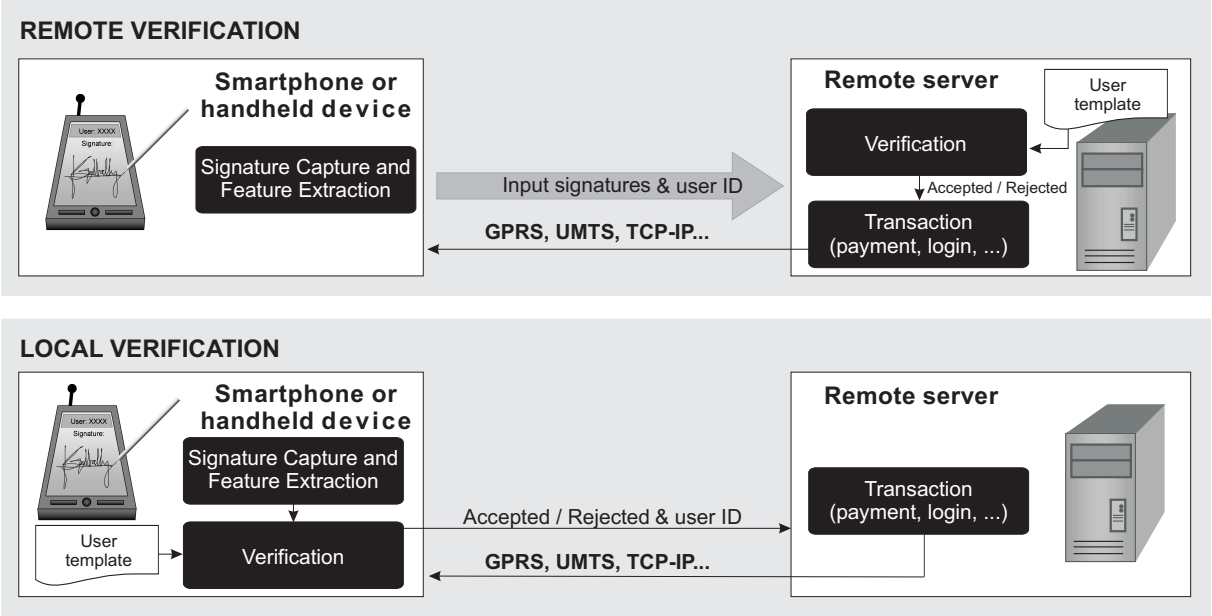


Figure 1.2: *Diagram of two possible architectures of a dynamic signature verification platform for handheld devices (The verification steps have been simplified).*

Paperless office: documents are electronically signed without printing them, providing verification of the signatures and ubiquitous access to them. This allows business process and workflow automation where signatures are needed.

In all these applications, the verification system can be either remote or local. Local verification systems perform the matching process in the handheld device, while remote systems send the input signatures over the network and the matching is performed on a remote server. A model of the two aforementioned architectures is presented in Fig. 1.2. Security must be ensured in both architectures. While in local systems, the user template and matcher must be secured in the handheld device, in remote systems, the transmitted signatures and verification system on the server side must be kept secure.

A key advantage for the deployment of such systems is that touchscreen mobile devices do not need any extra hardware for signature verification, as it is the case of fingerprint sensors or cameras for fingerprint and face verification systems respectively. Consequently, no extra costs exist and the system complexity does not increase.

1.4 Commercial Systems and Applications

Although signature verification systems have a small market share among biometric systems, there are many commercial systems available. Some examples of signature verification companies and products are the following:

- CIC provides secure signature verification applications oriented to the financial sector among other industries. It also provides an SDK (CIC).
- Crypto-Sign offers a PDA user authentication solution (Crypto-Sign).
- KeCrypt is aimed towards user authentication on workflows and payment systems. It provides both software and hardware (KeCrypt).
- Nemophila provides a signature verification software named Cyber-SIGN, mainly aimed towards client-server architectures (Nemophila).
- PDALok software is a PDA authentication solution (PdaLok).
- App Informatik Davos offers SignPlus, which is an offline signature verification and management system, primarily oriented to bank transactions (SignPlus).
- SoftPro solutions are focused on fraud detection and process securing (SOFTPRO).
- Topaz Systems provides software, SDKs and hardware for automatic signature verification (Topaz).
- Xyzmo offers signature verification solutions to a wide range of industries (Xyzmo).

1.5 Challenges of Signature Verification on Handheld Devices

Designers of signature verification systems must face many challenges. As has been previously stated, inter- and intra-variability represent two of the main difficulties when trying to reach a good verification performance, specially in the case of skilled forgeries.

Handheld devices such as smart phones or PDAs are affected by size and weight constraints due to their portable nature. While processing units, memory chips and battery components are nowadays experimenting higher levels of miniaturization and integration, the input (e.g. keyboard, touchscreen) and output (e.g. display) parts must have reasonable dimensions in order to keep their usability. Poor ergonomics and small input areas in mobile devices are two key factors that increase the variability during the signing process. Moreover, the unfamiliar signing surface may affect the signing process.

The touchscreen digitizing quality must also be taken into account. A typical digitizing tablet is based on an electromagnetic principle. The tablet has an embedded wire grid which acts as a transmitter. The pen (which is specifically designed for the tablet) acts as an antenna, which resonates and emits a signal that is captured by the tablet, allowing to detect its position with high accuracy. This allows the tablet to detect the pen movement even if it is not in contact with the tablet (in a reasonable range of proximity). On the other hand, touchscreens

of stylus-oriented handheld devices are based on a resistive principle. Two separated metallic layers are connected when the screen is pressed with a stylus. The position of the contact point can be accurately detected, but only when the surface is pressed.

Irregular sampling rates and sampling errors, which are common in mobile devices, may worsen the verification performance and must be addressed during the preprocessing steps. In these devices, only position signals are in general available. Pressure, pen-azimuth or other signals that may improve the verification performance (Muramatsu and Matsumoto, 2007), are not usually captured by touchscreens from handheld devices. The pen trajectory during pen-ups is recorded, in general, only in pen tablets. This information, which is invisible to forgers, is not available for PDAs due to the touchscreen technical limitations.

The interest in security on portable devices has raised in the last few years (Khokhar, 2006). Security must be a critical concern while designing a signature verification platform as a breach could give an attacker access to personal data or bank accounts. Gaining access to the matcher could allow an attacker to perform software attacks such as brute force or hill-climbing attacks (Galbally *et al.*, 2007b). The user template must be appropriately secured (Maiorana *et al.*, 2008) and encrypted (Freire-Santos *et al.*, 2006) as well as the communication channels over which signature information may be transmitted.

Chapter 2

Related Works and State of the Art

In this chapter, a summary of the research in dynamic signature verification is performed, presenting related works and available resources.

2.1 Architecture of a Dynamic Signature Verification System

Dynamic signature verification systems generally share a common architecture. The typical architecture of an automatic signature verification system is depicted in Fig. 2.1. In general, dynamic signature verification systems perform the following steps (Fierrez and Ortega-Garcia, 2007b):

1. **Data Acquisition**: Signature signals are captured using a digitizing tablet (WACOM) or touch-screen device (HYPERCOM) as a PDA or Tablet-PC. Special pens with sensors that capture forces and movement have also been proposed (Martens and Claesen, 1997). The signature signal is sampled and stored as discrete time series. While some digitizing tablets provide pressure or pen angle information, these are not commonly available in handheld devices. In all cases, the sampling rate is usually equal to or above 100 Hz. This is a reasonable rate, since it has been observed that the maximum frequencies of the signature time functions are approximately of 20 - 30 Hz (Plamondon and Lorette, 1989).

   After data acquisition, preprocessing steps are commonly performed. These include noise filtering, resampling, or interpolation of missing samples.

2. **Feature Extraction**: Two main approaches have been followed in this step: *feature-based* systems extract global features (e.g. signature duration, number of pen-ups, average velocity) from the signature in order to obtain a holistic feature vector (Lee *et al.*, 1996). On the other hand, *function-based* systems use the signature time functions (e.g. position, pressure) for verification. Traditionally, function-based approaches have yielded better

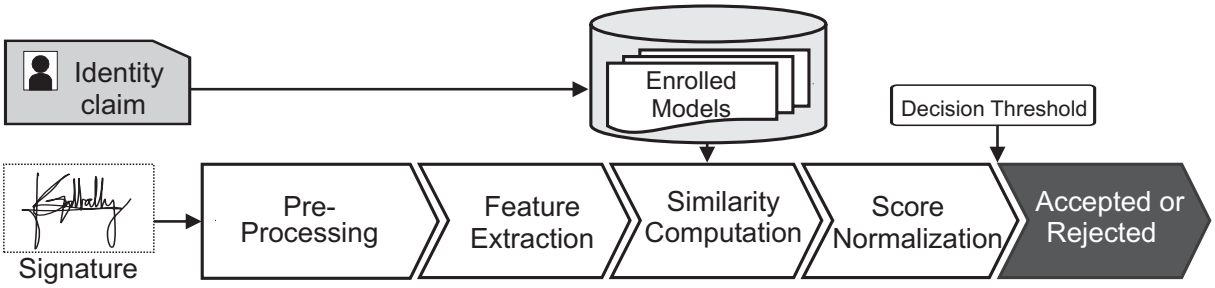


Figure 2.1: *Typical architecture of a signature verification system.*

results than feature-based ones (Fierrez-Aguilar *et al.*, 2005a; Kholmatov and Yanikoglu, 2005).

3. **Enrollment**: In *model-based* systems a statistical user model is computed using a training set of genuine signatures which is used for future comparisons in the matching step (Nanni and Lumini, 2005; Richiardi and Drygajlo, 2003). *Reference-based* systems store the features of each signature of the training set as templates. In the matching process the input signature is compared with each reference signature (Lei and Govindaraju, 2005).

4. **Similarity Computation**: This step involves *pre-alignment* if necessary and a *matching* process, which returns a *matching score*. In feature-based systems, statistical techniques like Mahalanobis distance, Parzen Windows or Neural Networks are used for matching (Nelson *et al.*, 1994). Function-based systems use other techniques like Hidden Markov Models - HMM (Dolfing *et al.*, 1998; Fierrez *et al.*, 2007; Van *et al.*, 2007), or Dynamic Time Warping - DTW (Kholmatov and Yanikoglu, 2005; Martens and Claesen, 1997; Sato and Kogure, 1982) to compare signature models.

5. **Score Normalization**: The matching score may be normalized to a given range. Score normalization is critical when combining scores from multiple classifiers or in multi-biometric systems (Ross *et al.*, 2006). More sophisticated techniques like target-dependent score normalization can lead to an improved system performance (Fierrez-Aguilar *et al.*, 2005b; Martinez-Diaz *et al.*, 2007b).

An input signature will be considered from the claimed user if its matching score exceeds a given threshold.

2.2 Feature-based Systems

Feature-based systems, also known as global systems, have been extensively studied (Fierrez-Aguilar *et al.*, 2005a; Lee *et al.*, 1996; Lei and Govindaraju, 2005; Richiardi *et al.*, 2005). In these

systems, a holistic vector is formed by features extracted from the whole signature, such as duration, average speed, number of pen-ups, etc. Despite the large amount of different global feature sets that have been proposed (a maximum of 100 features are considered in (Fierrez-Aguilar *et al.*, 2005a)), the usually low amount of available training data motivates the usage of feature selection techniques to reduce the feature vector size (due to the curse of dimensionality). Several feature selection techniques have been proposed (see Sect. 2.4), being the Sequential Forward Feature Selection (SFFS) (Pudil *et al.*, 1994) one of the best performing methods reported (Jain and Zongker, 1997). The matching phase is usually performed with statistical classifiers such as Gaussian Mixture Models (Martinez-Diaz *et al.*, 2007b), Parzen Windows (Martinez-Diaz *et al.*, 2007b), majority voting (Lee *et al.*, 1996), Mahalanobis distance (Galbally *et al.*, 2007b), etc.

2.3 Function-based Systems

Function-based systems are also known as local systems. Among these, signature verification systems using Dynamic Time Warping (DTW) (Kholmatov and Yanikoglu, 2005; Martens and Claesen, 1997; Sato and Kogure, 1982) or Hidden Markov Models (HMM) (Dolfing *et al.*, 1998; Fierrez *et al.*, 2007; Van *et al.*, 2007; Yang *et al.*, 1995) are the most popular approaches in signature verification. In these systems, the captured time functions (e.g. pen coordinates, pressure, etc.) are used to model each user signature. Additionally, the use of pen orientation features such as azimuth or altitude has been reported to provide good results (Muramatsu and Matsumoto, 2007), although it has been discussed by other authors (Lei and Govindaraju, 2005). Fusion of the feature- and function-based approaches has been reported to provide better performance than the individual systems (Fierrez-Aguilar *et al.*, 2005a).

2.3.1 Dynamic Time Warping

Dynamic Time Warping (DTW) is an application of Dynamic Programming to the problem of matching time sequences. Yasuhara and Oka (1977) were the first to report its suitability for dynamic signature verification, by using the algorithm to match time functions extracted from digitized signature signals. Their approach was an adaptation of the original algorithm proposed by Sakoe and Chiba (1978) in the field of speech recognition. The goal of DTW is to find an elastic match among samples of a pair of sequences X and Y that minimize a given distance measure. The algorithm may be defined as follows (Sakoe and Chiba, 1978). Let's define two sequences

$$\begin{aligned} X &= x_1, x_2, ..., x_i, ..., x_I \\ Y &= y_1, y_2, ..., y_j, ..., y_J \end{aligned} \tag{2.1}$$

and a distance measure as

$$d(i,j) = \|x_i - y_j\| \tag{2.2}$$

between sequence samples. A warping path can be defined as

$$C = c_1, c_2, ..., c_k, ..., c_K \tag{2.3}$$

where each c_k represents a correspondence (i,j) between samples of X and Y. The initial condition of the algorithm is set to

$$g_1 = g(1,1) = d(1,1) \cdot w(1) \tag{2.4}$$

where g_k represents the accumulated distance after k steps and $w(k)$ is a weighting factor that must be defined. For each iteration, g_k is computed as

$$g_k = g(i,j) = \min_{c_{k-1}} [g_{k-1} + d(c_k) \cdot w(k)] \tag{2.5}$$

until the I'th and J'th sample of both sequences respectively is reached. The resulting normalized distance is

$$D(X,Y) = \frac{g_K}{\sum_{k=1}^{K} w(k)} \tag{2.6}$$

where $\sum w(k)$ compensates the effect of the length of the sequences.

The weighting factors w_k are defined in order to restrict which correspondences among samples of both sequences are allowed. In Fig. 2.2.a, a possible definition of w_k is depicted. In this case, only three transitions are allowed in the computation of g_k. Consequently, Eq. (2.5) becomes

$$g_k = g(i,j) = \min \begin{bmatrix} g(i,j-1) + d(i,j) \\ g(i-1,j-1) + 2d(i,j) \\ g(i-1,j) + d(i,j) \end{bmatrix} \tag{2.7}$$

which is one of the most common implementations found in the literature. In Fig. 2.2.b, an example of point correspondences between two signatures is depicted to visually show how the results of the elastic alignment.

The algorithm has been further refined for signature verification by many authors (Kholmatov and Yanikoglu, 2005; Martens and Claesen, 1997; Sato and Kogure, 1982). Moreover, the implementation by Kholmatov and Yanikoglu (2005) won the Signature Verification Competition 2004 (Yeung *et al.*, 2004). Although the DTW algorithm has been replaced by more powerful ones such as HMMs or SVMs for speech applications, it remains as a highly effective

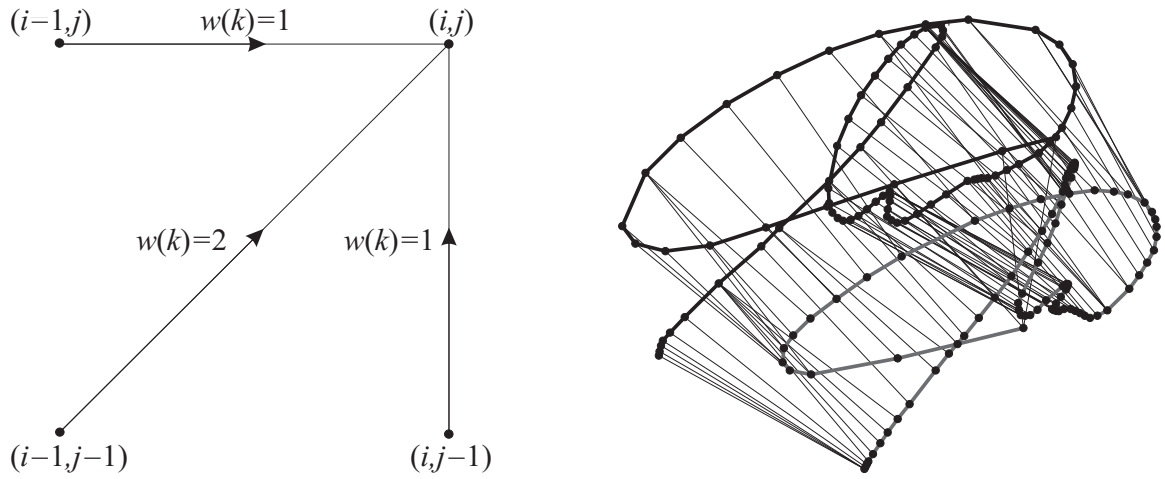


Figure 2.2: *(a) Example of symmetrical weighting factors $w(k)$ for Dynamic Time Warping. (b) Example of point-to-point correspondences between two genuine signatures obtained using DTW.*

tool for signature verification as it is best suited for small amounts of training data, which the common case in signature verification..

2.3.2 Hidden Markov Models

Hidden Markov Models (HMM) have been widely used by the speech recognition community (Rabiner, 1989) as well as in many handwriting recognition applications (Dolfing, 1998). Several approaches using HMMs for dynamic signature verification have been proposed in the last years (Dolfing *et al.*, 1998; Fierrez *et al.*, 2007; Muramatsu and Matsumoto, 2003; Van *et al.*, 2007; Yang *et al.*, 1995). An HMM represents a double stochastic process, governed by an underlying Markov chain, with a finite number of states and random function set that generate symbols or observations each of which is associated with one state (Yang *et al.*, 1995). Observations are modeled with GMMs in most speech and handwriting recognition applications. GMMs, which can be considered a single-state HMM, have been also successfully used for signature verification (Richiardi and Drygajlo, 2003).

The basic structure of an HMM using GMMs to model observations is defined by the following elements:

- Number of hidden states N.
- Number of Gaussian Mixtures per state M.
- Probability transition matrix $\mathbf{A} = \{a_{ij}\}$, which contains the probabilities of jumping from one state to another or staying on the same state.

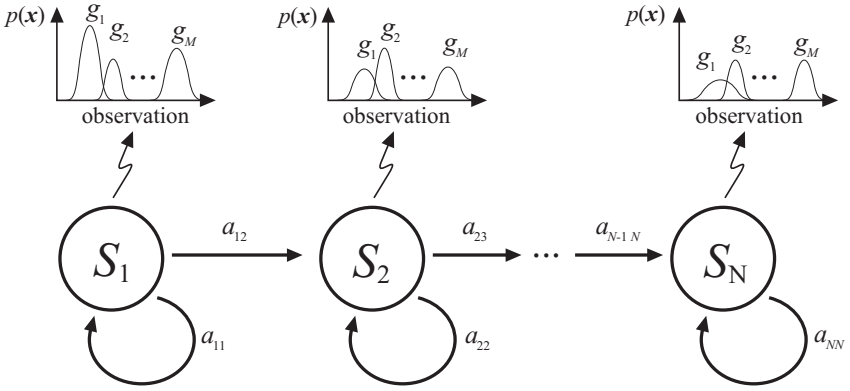


Figure 2.3: *Graphical representation of a left-to-right N-state HMM, with M-component GMMs representing observations and no skips between states.*

In Fig. 2.3, an example of a possible HMM configuration is shown. Hidden Markov Models are usually trained in two steps. First, state transition probabilities and observation statistical models are estimated using a Maximum Likelihood algorithm. After this, a re-estimation step is carried out using the Baum-Welch algorithm. A detailed description of the training process is given by Rabiner (1989).

Within HMM-based dynamic signature verification, *regional* and *local* approaches have been proposed. In regional approaches, the extracted time functions are further segmented and converted into a sequence of feature vectors or observations, each one representing regional properties of the signature signal (Dolfing *et al.*, 1998; Kashi *et al.*, 1997; Yang *et al.*, 1995). Some examples of segmentation boundaries are null vertical velocity points (Dolfing *et al.*, 1998) or changes in the quantized trajectory direction (Yang *et al.*, 1995). On the other hand, local approaches directly use the time functions as observation sequences for the signature modeling (Fierrez *et al.*, 2007; Richiardi and Drygajlo, 2003; Van *et al.*, 2007).

Finding a reliable and robust model structure for dynamic signature verification is not a trivial task. While too simple HMMs may not allow to model properly the user signatures, too complex models may not be able to model future realizations due to overfitting. On the other hand, as simple models have less parameters to be estimated, their estimation may be more robust than for complex models. Two main parameters are commonly considered while selecting an optimal model structure: the number of states and the number of Gaussian mixtures per state (Fierrez *et al.*, 2007). Most of the proposed systems consider a left-ro-right configuration without skips between states, also known as Bakis topology (see Fig. 2.3).

2.4 Feature Selection

Due to the curse of dimensionality (Theodoridis and Koutroumbas, 2006), the performance of a statistical classifier is degraded if the available training data is too small compared to the number of dimensions of the feature vector (Jain and Zongker, 1997). This is usually the case in signature verification, where the average length of a digitized signature is of a few hundreds of samples and the available number of training signatures is relatively small (in practical applications between 3 and 5). The amount of training signatures is mostly conditioned by the willingness of the users to provide many samples during enrollment. Nevertheless, when signatures are captured during only one unique session, their variability is small in general, leading to a poorly trained model.

Feature selection techniques try to reduce the dimensionality of the feature vectors while optimizing the verification accuracy. Their goal is to find the optimum combination of features according to a given optimization criterion. Ideally, given a feature vector of F dimensions, all the possible combinations from 1 to F features should be tested in order to find the optimal combination. Unfortunately, this is not feasible due to the high amount of combinations that have to be tested, which is

$$\sum_{i=1}^{F} \binom{F}{i}.$$

A critical step when performing feature selection is the choice of the optimization criterion. Two main alternatives can be taken: *filter* and *wrapper* methods (Theodoridis and Koutroumbas, 2006). In the former, the optimal feature subset is selected according to intrinsic properties of the training data such as statistical properties. In the latter, the result of the classification problem under consideration is used as the criterion to be optimized. A reasonable choice for a signature verification system is a wrapper method in which the verification performance in terms of the EER is set as the optimization criterion. Wrapper methods require in general more computational resources, as the evaluation of the optimization criterion (e.g. the verification decision) is commonly more complex than the computation of statistical properties of the training data.

Feature selection has been applied to signature verification from several perspectives. Lee *et al.* (1996) propose a method for global features which ranks the discriminative power of each feature for each specific user, based on the distance between the user signatures and the rest of users. They select as an optimal feature vector the one that contains the features that are most commonly ranked among the most discriminative over all the users in the database. Fierrez-Aguilar *et al.* (2005a) perform feature ranking based on their Mahalanobis distance between signatures from different users. The optimal feature vector is then selected by iteratively adding individual features in the order they were ranked and selecting the best performing vector in

terms of the system EER. Richiardi *et al.* (2005) propose a distance measure based on the Fisher Discriminant Ratio and use it to perform Sequential Forward Floating Search Selection (SFFS), which is summarized in this section. Galbally *et al.* (2007a) perform feature selection by using Genetic Algorithms and setting the system EER as the optimization criterion.

2.4.1 Feature Selection Algorithms

Several feature selection techniques have been proposed in the literature aimed towards reducing the number of feature combinations that have to be tested. Unfortunately, all of them are only able to find suboptimal solutions. A notable exception is the Branch and Bound algorithm, which is however only applicable when the optimization criterion increases monotonically with the feature subset size. While some of the algorithms are deterministic and always lead to the same suboptimal solution, other algorithms may produce different suboptimal solutions in each execution (Jain and Zongker, 1997). The most popular techniques are summarized next.

2.4.1.1 Scalar Feature Selection

Features are considered individually using this procedure. Each feature is ranked in terms of its class separability using a predefined criterion C, such as the system EER or any distance measure. Then, the N top ranked features in terms of C are selected as the optimal N-dimensional feature vector. This method has the advantage of being computationally simple. Nevertheless, it does not take into account the possible correlations among features. Some techniques to deal with this problem have been proposed in the literature (Theodoridis and Koutroumbas, 2006). This approach is taken in (Fierrez-Aguilar *et al.*, 2005a).

2.4.1.2 Sequential Forward/Backward Selection

In Sequential Forward Selection, given F available features we start by selecting the most discriminative feature x_i. Then, all the possible combinations between this feature and any other feature x_f are computed and the best combination $\{x_i, x_j\}$ is selected. The algorithm continues by iteratively adding features in this manner until the desired feature vector size is reached. The Sequential Backward Selection is similar to this approach but instead of starting with a single feature it starts with all the F features and one feature is removed at a time.

2.4.1.3 Floating Search

Pudil *et al.* (1994) proposed a feature selection algorithm that overcomes some of the limitations of the ones presented above. Namely, when a feature is selected by the previous methods (or discarded, in the backward case), it can no longer be discarded (or selected, in the backward case). This is known as the nesting effect. As with Sequential Selection, both a forward and

a backward approach exist. We focus on the forward method, referred to as Sequential Forward Floating Search (SFFS). The algorithm can be summarized as follows (Theodoridis and Koutroumbas, 2006).

Let's consider a set of F features, from which we wish to find the best performing subset of N features, $N \leq F$ in terms of a given criterion C. Let $X_n = \{x_1, x_2, ..., x_n\}$ be the best combination of n features and Y_{F-n} the set of remaining $F - n$ features. In the algorithm, we store the best sets of lower dimensions $X_1, X_2, ..., X_{n-1}$. The following steps are performed until a loop with a stable set X_n is obtained.

1. *Inclusion*
   Choose the element x_{n+1} from Y_{F-n} which, added to X_n produces the best value of the optimization criterion C. Then, $X_{n+1} = \{X_n, x_{n+1}\}$.

2. *Test*

   (a) Find the feature x_r that has the least negative (or most positive) effect on the criterion C when it is removed from X_{n+1}.

   (b) If $r = n + 1$, change n for $n + 1$ and go to step 1.

   (c) If $r \neq n + 1$ and $C(X_{n+1} - \{x_r\}) < C(X_n)$ go to step 1, that is, if removal of any feature does not improve the criterion on the previously selected set X_n, no further backward search is performed.

3. *Exclusion*

   (a) Remove x_r to get $X'_n = X_{n+1} - \{x_r\}$.

   (b) Find the feature x_s that has the least negative effect on the criterion C when it is removed from X'_n.

   (c) If $C(X'_n - \{x_s\}) < C(X_{n-1})$ then $X_n = X'_n$ and go to step 1, that is, if removal of another feature does not improve the criterion on the previously selected set X_n, no further backward search is performed.

   (d) Remove x_s by putting $X'_{n-1} = X'_n - \{x_s\}$ and $n = n - 1$.

   (e) Go to step 3.a.

Note that some specific conditions on the first steps have not been considered in order to simplify the algorithm description. The backward algorithm is equivalent to the one explained but removing instead of adding features.

Other algorithms for feature selection include Neural Networks and Genetic Algorithms (Gallbally *et al.*, 2007a), although the latter produce variable suboptimal results among different

executions. Jain and Zongker (1997) performed an exhaustive comparison of several feature selection algorithms and studied the impact of small training sets on them. The SFFS proved to be highly effective, obtaining results close to the optimal set selected by the Branch and Bound algorithm.

2.5 Dynamic Signature Databases

Publicly available databases allow researchers to perform an objective comparison of their verification algorithms. Until the last few years, much research had been carried out using private databases, as no large public ones were available. This does not allow reliable performance comparisons of different algorithms, which may have been tuned to a specific database. Moreover, the usage of small datasets reduces the statistical relevance of experiments. Privacy and legal issues have also played a relevant role in the lack of public signature datasets.

The variation of signatures among different cultures must also be taken into account. As an example, in Europe signatures are usually formed by a fast writing followed by a flourish while in North America they usually correspond to the signers name with no flourish. On the other hand, signatures in Asia are commonly formed by asian characters, which are composed of a larger number of short strokes compared to European or North American signatures.

While some authors have made public the databases used for their research (Munich and Perona, 2003), most current dynamic signature databases are collected by the joint effort of different research institutions. In the following, a brief description of the most relevant available databases, in chronological order, is given.

PHILIPS Database. Signatures from 51 users were captured using a digitizing tablet at a sampling rate of 200 Hz (Dolfing *et al.*, 1998). The following signals were captured: pen-coordinates, pen-pressure, and the pen-tilt, which is composed by the two angles resulting from the projection of the pen in the (x, z) and (y, z) planes.

Each user contributed 30 genuine signatures, leading to 1530 genuine signatures. Three types of forgeries are present in the database: 1470 over-the-shoulder forgeries, 1530 home-improved and 240 professional forgeries. Over-the-shoulder forgeries were produced by letting the forger observe the signing process. Home-improved forgeries were produced by giving to the forgers the signature static image and letting them to practice at home (see Fig. 2.4). Finally, professional forgeries were performed by forensic document examiners.

MCYT Signature Subcorpus. The MCYT bimodal database is comprised of signatures and fingerprints from 330 individuals (Ortega-Garcia *et al.*, 2003). Signatures were acquired using a WACOM Intuos A6 tablet with a sampling frequency of 100 Hz. The capture area

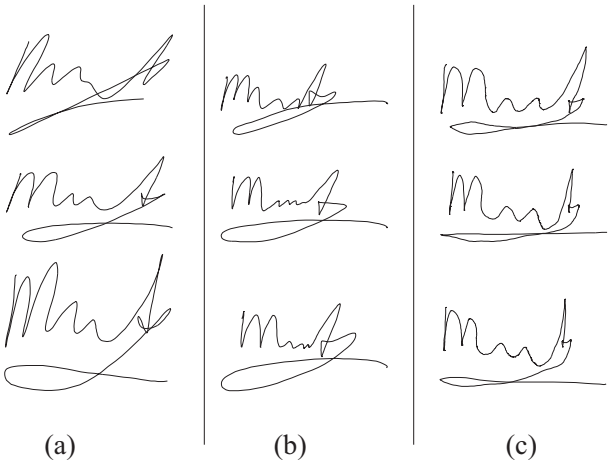


Figure 2.4: *Examples of signatures for a particular subject of the PHILIPS Database. (a) Genuine signatures, (b) over-the-shoulder forgeries, and (c) home improved forgeries. Adapted from (Dolfing et al., 1998).*

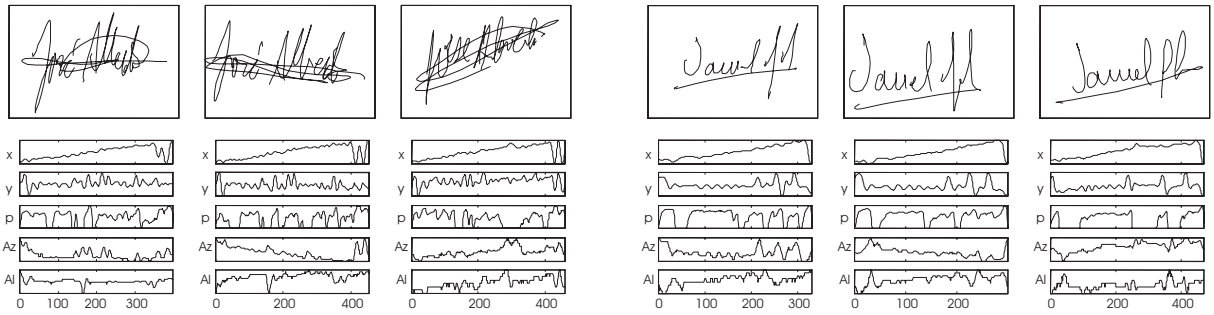


Figure 2.5: *MCYT example signatures and associated functions for two different subjects. One genuine signature (left) and two forgeries (right columns) are presented for each user. Adapted from (Fierrez et al., 2007).*

was divided in frames for acquisition of 37.5 mm (width) × 17.5 mm (height). The following time sequences are captured: position coordinates, pressure, azimuth angle and altitude angle. Example signatures and their associated functions are shown in Fig. 2.5.

There are 25 genuine signatures and 25 forgeries per user. Signatures were captured in groups of 5. First, 5 genuine signatures, then 5 skilled forgeries from another user, repeating this until 25 signatures from each type were performed. Each user provided 5 forgeries for the 5 previous users in the database. As the user is forced to concentrate on different tasks between each group of genuine signatures, the variability between groups is expected to be higher than the one within the same group.

BIOMET Signature Subcorpus. This signature subcorpus is part of the BIOMET multimodal database (Garcia-Salicetti *et al.*, 2003). The signatures were captured using a WACOM Intuos2 A6 Pen-tablet and an ink pen with a sampling rate of 100 Hz. The pen coordinates,

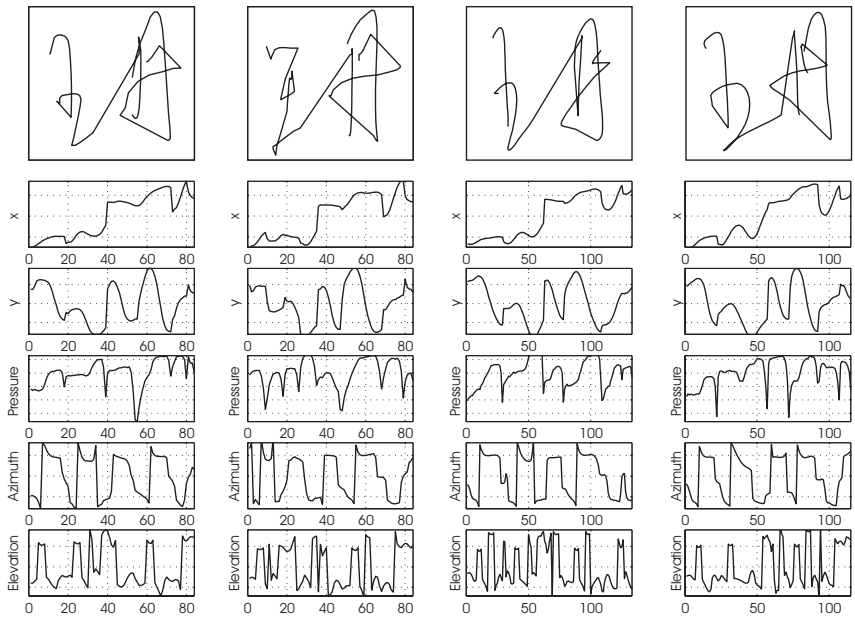
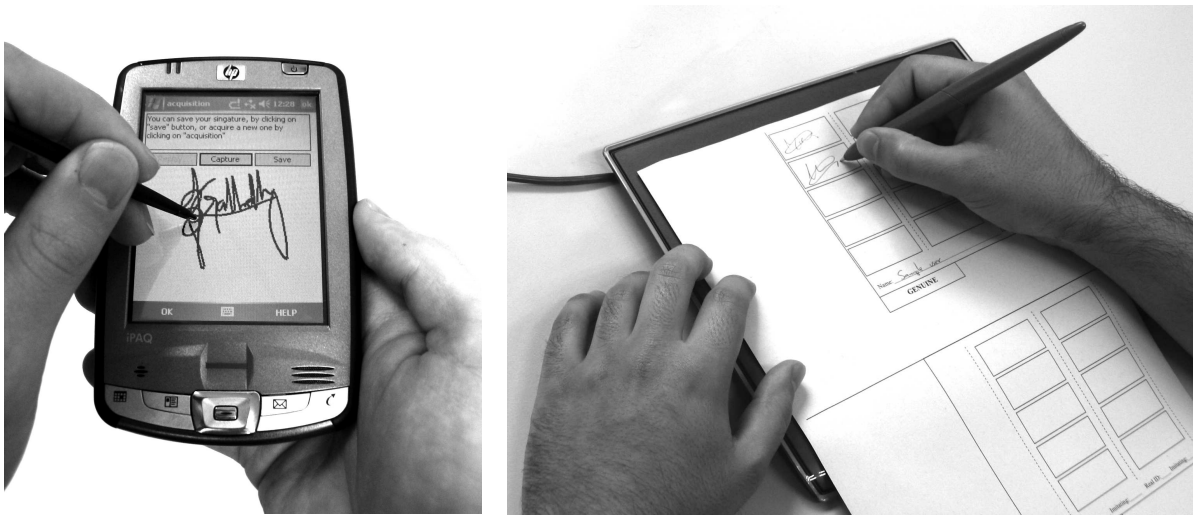


Figure 2.6: *SVC 2004 example signatures and associated functions extracted by the pen tablet. For a particular subject, two genuine signatures (left columns) and two forgeries (right columns) are presented. Adapted from (Fierrez* et al.*, 2007).*

pen-pressure, azimuth and altitude signals were captured. The database contains data from 84 users, with 15 genuine signatures and 12 forgeries per user. Signatures were captured in two sessions separated by 5 months. In the first session, 5 genuine signatures and 6 forgeries were acquired. The remaining 10 genuine signatures and 6 forgeries were captured in the second session. Forgeries are performed by 4 different users (3 forgeries each). This database contains 2201 signatures, since not all users have complete data: 8 genuine signatures and 54 forgeries are missing.

SVC 2004 Database. Two development databases were released prior to the Signature Verification Competition (SVC) 2004 (Yeung *et al.*, 2004). They were captured using a WACOM digitizing tablet and a Grip Pen. Due to privacy issues, users were advised to use invented signatures as genuine ones. The two databases differ in the available data, and correspond to the two tasks defined in the competition. One contains only coordinate information while the other provides also pressure and pen orientation signals. Each database contains 40 users, with 20 genuine signatures and 20 forgeries per user acquired in two sessions. Both occidental and asian signatures are present in the databases. Examples of signatures from this database are shown in Fig. 2.6.

BioSecure Signature Subcorpus DS2 - Access Control Scenario. This database was captured under the BioSecure Network of Excellence (Biosecure Network of Excellence, 2007)



(a) (b)

Figure 2.7: *(a) PDA signature capture process in the BIOSECURE DS3 - Mobile Scenario dataset. (b) Pen-tablet capture process in the BIOSECURE DS2 - Access Control Scenario dataset.*

and is not yet available. It is part of the BioSecure multimodal database (Data Set 2) and consists of 667 users, although the completeness of the data for each user is still not confirmed. It was acquired at seven different sites in Europe. Acquisition was carried out using a WACOM Intuos3 A6 digitizer at 100 Hz following a procedure similar to the one conducted in MCYT (Ortega-Garcia *et al.*, 2003). The pen coordinates, pressure, azimuth and altitude signals are available.

Signatures were captured in two sessions and in blocks of 5. During each session, users were asked to perform 3 sets of 5 genuine signatures, and 5 forgeries between each set. Each user performed 5 forgeries for the previous 4 users in the database. The users had visual access to the dynamics of the signing process of the signatures they had to forge. Thus, 30 genuine signatures and 20 forgeries are available for each user. An example of the signature capture process of this database including the paper template that was used is depicted in Fig. 2.7.(b).

BioSecure Signature Subcorpus DS3 - Mobile Scenario. The BioSecure Signature Subcorpus DS3 was acquired under the same framework than the Access Control Scenario but on a mobile scenario (GET-INT, 2007). It was acquired in 8 different sites in Europe and is not yet available (Alonso-Fernandez *et al.*, 2008). It is the first multi-session database captured on a PDA. An HP iPAQ hx2790 with a sampling frequency of 100 Hz was used as capture device. Only the pen coordinates and time stamps are available. Users were asked to sign while standing and holding the PDA in one hand. This was done to emulate realistic operating conditions. The acquisition protocol was the same than for the Access Control Scenario Signature Subcorpus,

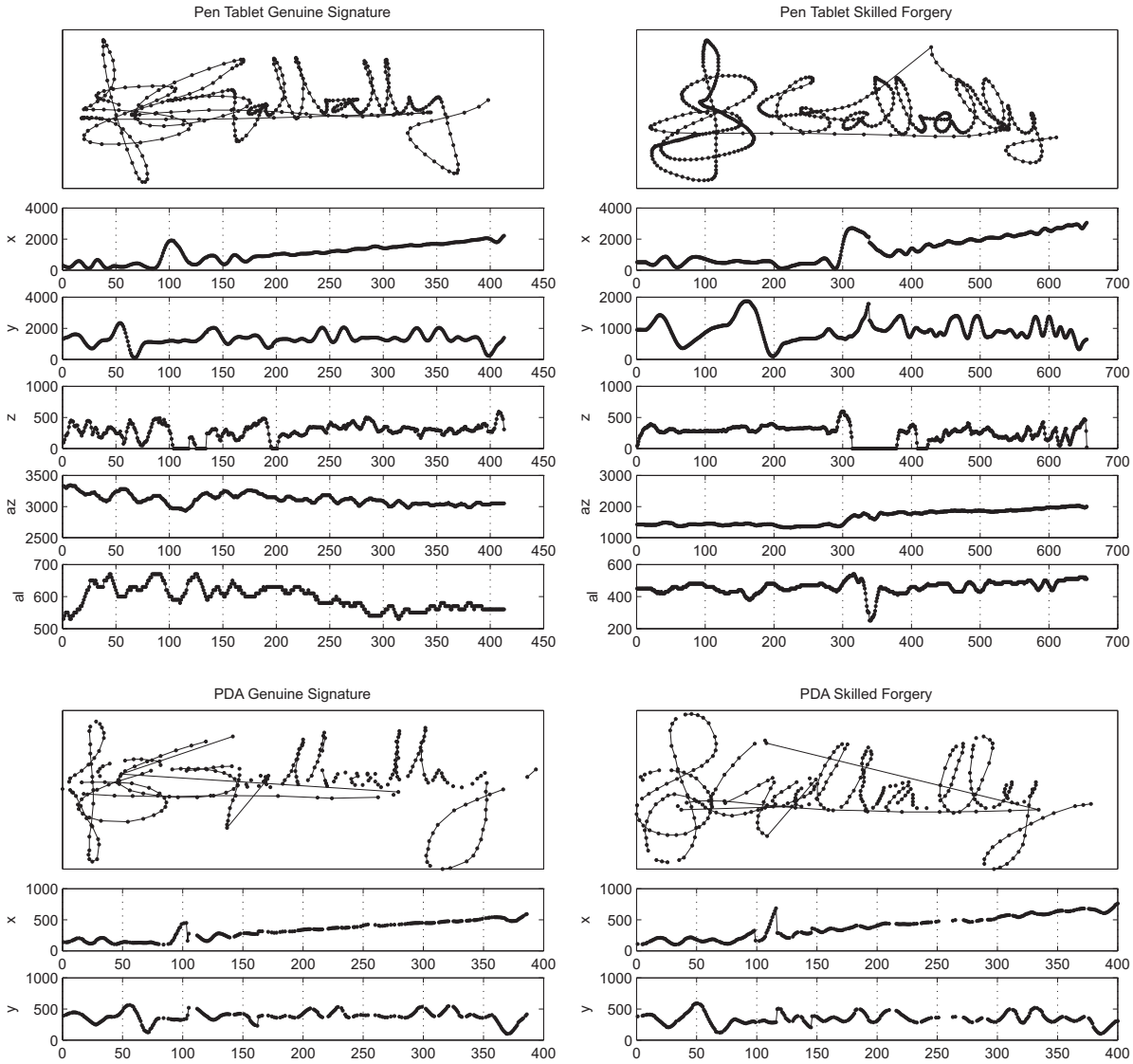


Figure 2.8: *Examples of signatures and associated signals from the BioSecure Multimodal Database DS2 and DS3 signature subcorpora captured using a pen tablet (top) and a PDA (bottom), respectively.*

in which signature data was captured using a pen tablet. An average of two months was left between each session. Forgeries for each user are performed by 4 different users (5 forgeries each) in a "worst case" scenario, where each forger has access to the dynamics of the genuine signature in the PDA screen and a tracker tool allowing to see the original strokes. An example of the capture process of this database can be seen in Fig. 2.7.(a). Examples of signatures from the BioSecure Signature subcorpora DS2 and DS3 are presented in Fig. 2.8. Signatures captured with the PDA present missing samples (i.e. sampling errors) due to the PDA touchscreen acquisition process.

Other databases found in the literature include the SUSIG Database (Kholmatov and Yanikoglu, 2008) and the MyIDea signature subcorpus (Dumas *et al.*, 2005).

Chapter 3

Signature Verification Systems

In the present chapter, the automatic signature verification systems used in this work are described. A function-based system (referred to as local) and a feature-based system (referred to as global) are considered. The local verification system is based on the one presented by Fierrez and Ortega-Garcia (2007a), to which some features extracted from related works have been added. The global system is based on the one presented in (Fierrez-Aguilar *et al.*, 2005a).

3.1 Global Signature Verification System

This feature-based signature verification system considers a set of 100 global features, extracted from each signature. The feature set is an extension of other sets presented in previous works in the literature (Lee *et al.*, 1996; Nelson and Kishon, 1991; Nelson *et al.*, 1994). A complete description of the feature set is shown in Table 3.1. These 100 features can be divided in four categories corresponding to the following magnitudes (the numbering is the same used in Fierrez-Aguilar *et al.* (2005a)):

- **Time** (25 features), related to signature duration, or timing of events such as pen-ups or local maxima: 1, 13, 22, 32, 38, 40-42, 50, 52, 58-60, 62, 64, 68, 79, 81-82, 87-90, 94, 100.
- **Speed and Acceleration** (25 features), from the first and second order time derivatives of the position time functions, like average speed or maximum speed: 4-6, 9-11, 14, 23, 26, 29, 31, 33, 39, 44-45, 48, 69, 74, 76, 80, 83, 85, 91-92, 96.
- **Direction** (18 features), extracted from the path trajectory like the starting direction or mean direction between pen-ups: 34, 51, 56-57, 61, 63, 66, 71-73, 77-78, 84, 93, 95, 97-99.
- **Geometry** (32 features), associated to the strokes or signature aspect-ratio: 2, 3, 7-8, 12, 15-21, 24-25, 27-28, 30, 35-37, 43, 46-47, 49, 53-55, 65, 67, 70, 75, 86.

Features are normalized into the range $(0, 1)$ using tanh-estimators (Jain *et al.*, 2005). The Mahalanobis distance is used to compare a signature with a claimed user model. This distance

Table 3.1: *Set of global features. Table adapted from (Fierrez-Aguilar et al., 2005a). T denotes time interval, t denotes time instant, N denotes number of events, and θ denotes angle. Note that some symbols are defined in different features of the table (e.g. Δ in feature 7 is defined in feature 15)*

#	Time related feature	#	Direction related feature
#	Speed and Acceleration related feature	#	Geometry related feature

Ranking	Feature Description	Ranking	Feature Description
1	signature total duration T_s	2	N(pen-ups)
3	N(sign changes of dx/dt and dy/dt)	4	average jerk $\bar{\jmath}$
5	standard deviation of a_y	6	standard deviation of v_y
7	(standard deviation of y)$/\Delta_y$	8	N(local maxima in x)
9	standard deviation of a_x	10	standard deviation of v_x
11	j_{rms}	12	N(local maxima in y)
13	t(2nd pen-down)$/T_s$	14	(average velocity $\bar{v}$)$/v_{x,\max}$
15	$\frac{A_{\min}=(y_{\max}-y_{\min})(x_{\max}-x_{\min})}{(\Delta_x=\sum_{i=1}^{\text{pen-downs}}(x_{\max\,\vert i}-x_{\min\,\vert i}))\Delta_y}$	16	$(x_{\text{last pen-up}}-x_{\max})/\Delta_x$
17	$(x_{\text{1st pen-down}}-x_{\min})/\Delta_x$	18	$(y_{\text{last pen-up}}-y_{\min})/\Delta_y$
19	$(y_{\text{1st pen-down}}-y_{\min})/\Delta_y$	20	$(T_w\bar{v})/(y_{\max}-y_{\min})$
21	$(T_w\bar{v})/(x_{\max}-x_{\min})$	22	(pen-down duration T_w)$/T_s$
23	$\bar{v}/v_{y,\max}$	24	$(y_{\text{last pen-up}}-y_{\max})/\Delta_y$
25	$\frac{T((dy/dt)/(dx/dt)>0)}{T((dy/dt)/(dx/dt)<0)}$	26	$\bar{v}/v_{\max}$
27	$(y_{\text{1st pen-down}}-y_{\max})/\Delta_y$	28	$(x_{\text{last pen-up}}-x_{\min})/\Delta_x$
29	(velocity rms v)$/v_{\max}$	30	$\frac{(x_{\max}-x_{\min})\Delta_y}{(y_{\max}-y_{\min})\Delta_x}$
31	(velocity correlation $v_{x,y}$)$/v_{\max}^2$	32	$T(v_y>0\vert\text{pen-up})/T_w$
33	$N(v_x=0)$	34	direction histogram s_1
35	$(y_{\text{2nd local max}}-y_{\text{1st pen-down}})/\Delta_y$	36	$(x_{\max}-x_{\min})/x_{\text{acquisition range}}$
37	$(x_{\text{1st pen-down}}-x_{\max})/\Delta_x$	38	$T(\text{curvature}>\text{Threshold}_{\text{curv}})/T_w$
39	(integrated abs. centr. acc. a_{Ic})$/a_{\max}$	40	$T(v_x>0)/T_w$
41	$T(v_x<0\vert\text{pen-up})/T_w$	42	$T(v_x>0\vert\text{pen-up})/T_w$
43	$(x_{\text{3rd local max}}-x_{\text{1st pen-down}})/\Delta_x$	44	$N(v_y=0)$
45	(acceleration rms a)$/a_{\max}$	46	(standard deviation of x)$/\Delta_x$
47	$\frac{T((dx/dt)(dy/dt)>0)}{T((dx/dt)(dy/dt)<0)}$	48	(tangential acceleration rms a_t)$/a_{\max}$
49	$(x_{\text{2nd local max}}-x_{\text{1st pen-down}})/\Delta_x$	50	$T(v_y<0\vert\text{pen-up})/T_w$
51	direction histogram s_2	52	t(3rd pen-down)$/T_s$
53	(max distance between points)$/A_{\min}$	54	$(y_{\text{3rd local max}}-y_{\text{1st pen-down}})/\Delta_y$
55	$(\bar{x}-x_{\min})/\bar{x}$	56	direction histogram s_5
57	direction histogram s_3	58	$T(v_x<0)/T_w$
59	$T(v_y>0)/T_w$	60	$T(v_y<0)/T_w$
61	direction histogram s_8	62	(1st $t(v_{x,\min})$)$/T_w$
63	direction histogram s_6	64	T(1st pen-up)$/T_w$
65	spatial histogram t_4	66	direction histogram s_4
67	$(y_{\max}-y_{\min})/y_{\text{acquisition range}}$	68	(1st $t(v_{x,\max})$)$/T_w$
69	(centripetal acceleration rms a_c)$/a_{\max}$	70	spatial histogram t_1
71	θ(1st to 2nd pen-down)	72	θ(1st pen-down to 2nd pen-up)
73	direction histogram s_7	74	$t(j_{x,\max})/T_w$
75	spatial histogram t_2	76	$j_{x,\max}$
77	θ(1st pen-down to last pen-up)	78	θ(1st pen-down to 1st pen-up)
79	(1st $t(x_{\max})$)$/T_w$	80	$\bar{\jmath}_x$
81	T(2nd pen-up)$/T_w$	82	(1st $t(v_{\max})$)$/T_w$
83	$j_{y,\max}$	84	θ(2nd pen-down to 2nd pen-up)
85	$j_{\max}$	86	spatial histogram t_3
87	(1st $t(v_{y,\min})$)$/T_w$	88	(2nd $t(x_{\max})$)$/T_w$
89	(3rd $t(x_{\max})$)$/T_w$	90	(1st $t(v_{y,\max})$)$/T_w$
91	$t(j_{\max})/T_w$	92	$t(j_{y,\max})/T_w$
93	direction change histogram c_2	94	(3rd $t(y_{\max})$)$/T_w$
95	direction change histogram c_4	96	$\bar{\jmath}_y$
97	direction change histogram c_3	98	θ(initial direction)
99	θ(before last pen-up)	100	(2nd $t(y_{\max})$)$/T_w$

measure has the advantage of being relatively simple to compute and generic enough to provide a reasonable empirical estimate of the statistical class separability achieved by the features. In this manner, user models $C = (\boldsymbol{\mu}, \boldsymbol{\Sigma})$ are created from a training set of signatures, where $\boldsymbol{\Sigma}$ is a diagonal covariance matrix. Thus, a classifier is built where the matching score is obtained as the inverse of the Mahalanobis distance between the input signature feature vector $\mathbf{x}$ and the claimed user model C:

$$s(\mathbf{x}, C) = \left((\mathbf{x} - \boldsymbol{\mu})^T (\boldsymbol{\Sigma})^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right)^{-1/2}.$$

If the score $s(\mathbf{x}, C)$ is above a specific threshold, the signature is considered as genuine. On the contrary it is rejected by the system.

3.2 Local Signature Verification System

The system implemented in this work is based on the one described by Fierrez *et al.* (2007). It participated in the Signature Verification Competition 2004 (Yeung *et al.*, 2004), where it reached the first and second positions against random and skilled forgeries respectively.

The signals captured by the digitizer are used to extract a set of functions that model each signature. The original set of functions from (Fierrez *et al.*, 2007) has been extended in this work, adapting features from other contributions (Lei and Govindaraju, 2005; Richiardi *et al.*, 2005; Van *et al.*, 2007). In the original set, 7 functions were extracted from the raw signals, from which the first and second order derivatives were computed, leading to a 21-dimensional feature vector. The second order derivatives are discarded in this work as they proved to have a very low contribution in the verification performance (as corroborated in Richiardi *et al.* (2005)). In the present work, an extended set of 15 functions is proposed, plus 12 functions obtained from the first and second order derivatives of some of them. In Table 3.2 we present the resulting set of 27 functions. Examples of the extracted functions can be seen in Fig. 3.1.

As it will be further addressed, the original system assumes the availability of pressure and pen-inclination information, although this is not usually the case for signatures captured with a handheld device. In those cases, only 21 features can be extracted from the raw signals.

In our implementation, an initial step is added to the original HMM training scheme (Fierrez *et al.*, 2007), leading to the following stages: *i*) the global mean and covariance of the training signatures is assigned to all the mixtures, *ii*) k-means segmentation and Maximum Likelihood training is performed, *iii*) Baum-Welch re-estimation is carried out. The first step allows to have a trainable model for step *iii* (despite being inaccurate) in the case where step *ii* fails due to the large number of parameters to be estimated, or other computational problems.

Similarity scores are computed as the log-likelihood of the signature (using the Viterbi algorithm) divided by the total number of samples of the signature. No score alignment between users is applied (Fierrez-Aguilar *et al.*, 2005b). In order to keep scores between a reasonable range, normalized scores s_n between (0,1) are obtained as $s_n = \exp(s(\mathbf{x}, C)/30)$, where $\mathbf{x}$ and C represent respectively the input signature to verify and the enrolled model of the claimed identity.

Table 3.2: *Extended set of local features. The upper dot notation (e.g. $\dot{x}_n$) indicates time derivative. Features 3, 10, 15, 16, 17 and 18 are not available in the PDA scenario.*

#	Feature	Description
1	x-coordinate	x_n
2	y-coordinate	y_n
3	Pen-pressure	z_n
4	Path-tangent angle	$\theta_n = \arctan(\dot{y}_n/\dot{x}_n)$
5	Path velocity magnitude	$\upsilon_n = \sqrt{\dot{y}_n + \dot{x}_n}$
6	Log curvature radius	$\rho_n = \log(1/\kappa_n) = \log(\upsilon_n/\dot{\theta}_n)$, where κ_n is the curvature of the position trajectory
7	Total acceleration magnitude	$a_n = \sqrt{t_n^2 + c_n^2} = \sqrt{\dot{\upsilon}_n^2 + \upsilon_n^2\theta_n^2}$, where t_n and c_n are respectively the tangential and centripetal acceleration components of the pen motion.
8-14	First-order derivative of features 1-7	$\dot{x}_n, \dot{y}_n, \dot{z}_n, \dot{\theta}_n, \dot{\upsilon}_n, \dot{\rho}_n, \dot{a}_n$
15	Pen azimuth	γ_n
16	Pen altitude	ϕ_n
17-18	First-order derivative of features 15-16	$\dot{\gamma}_n, \dot{\phi}_n$
19-20	Second-order derivative of features 1-2	$\ddot{x}_n, \ddot{y}_n$
21	Ratio of the minimum over the maximum speed over a window of 5 samples	$\upsilon_n^r = \min\{\upsilon_{n-4}, ..., \upsilon_n\}/\max\{\upsilon_{n-4}, ..., \upsilon_n\}$
22-23	Angle of consecutive samples and first order difference	$\alpha_n = \arctan(y_n - y_{n-1}/x_n - x_{n-1})$ $\dot{\alpha}_n$
24	Sine	$s_n = \sin(\alpha_n)$
25	Cosine	$c_n = \cos(\alpha_n)$
26	Stroke length to width ratio over a window of 5 samples	$r_n^5 = \frac{\sum_{k=n-4}^{k=n}\sqrt{(x_k-x_{k-1})^2+(y_k-y_{k-1})^2}}{\max\{x_{n-4},...,x_n\}-\min\{x_{n-4},...,x_n\}}$
27	Stroke length to width ratio over a window of 7 samples	$r_n^7 = \frac{\sum_{k=n-6}^{k=n}\sqrt{(x_k-x_{k-1})^2+(y_k-y_{k-1})^2}}{\max\{x_{n-6},...,x_n\}-\min\{x_{n-6},...,x_n\}}$

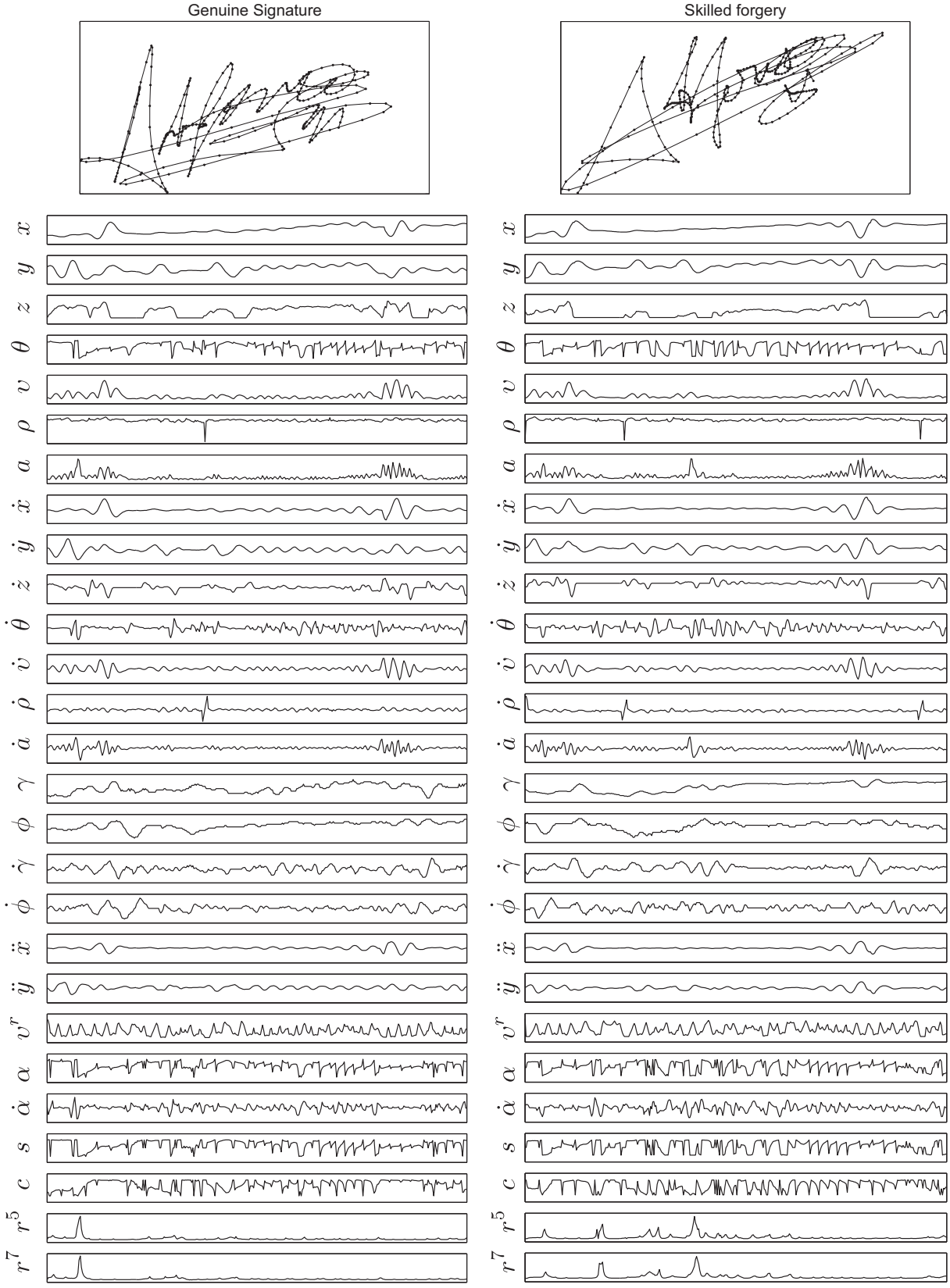


Figure 3.1: *Examples of functions from the 27-feature extended set defined in Table 3.2 for a genuine signature (left) and a skilled forgery (right) of a particular subject from the BIOSECURE DS2 Database.*

Chapter 4

Experiments

The experimental framework and results of this thesis are covered in this chapter. We have defined a protocol in which the particularities of signature verification for handheld devices and the pen tablet scenario are studied from a statistical and an applied point of view. Experiments are carried out using signatures from the same set of users acquired in the two scenarios considered, allowing to perform a fair comparison between them.

4.1 Database

A subset of the PDA and pen tablet signature corpus of the BioSecure Multimodal Biometric Database (Alonso-Fernandez *et al.*, 2008) is used for experiments. It consists of 120 users, with 20 genuine signatures and 20 skilled forgeries per user and acquisition device (PDA and pen tablet). The genuine signatures were acquired in two different sessions separated by an average period of two months, being 5 signatures from the first session and the remaining 15 from the second session. In each session, signatures were performed in blocks of 5, leaving a gap of some minutes between each block. Signatures were captured with an HP iPAQ 2790 PDA while the user was standing and holding the PDA with one hand; whereas for the pen tablet case they were captured with a Wacom Intuos A6 tablet while the user was sitting, using a pen on a paper placed over the tablet (see Fig. 2.7). This emulates real operating conditions.

Only the x and y position signals and the sample timestamps are captured by the PDA, while pressure information and pen orientation (azimuth and altitude angles) is also provided by the pen tablet. Skilled forgeries for each user were performed by 4 different users (5 forgeries each) in a "worst case" scenario: each forger had visual access to the dynamics of the genuine signature and a tracker tool allowing to see the original strokes in both scenarios. In the PDA scenario, the original strokes were shown in the PDA screen, where some users were even allowed to sign while seeing them. For the pen tablet scenario, the strokes were presented in a computer screen.

An example of the captured signatures, their associated signals and the distribution of two features on both scenarios is shown in Fig. 4.1. It can be seen that signatures captured with

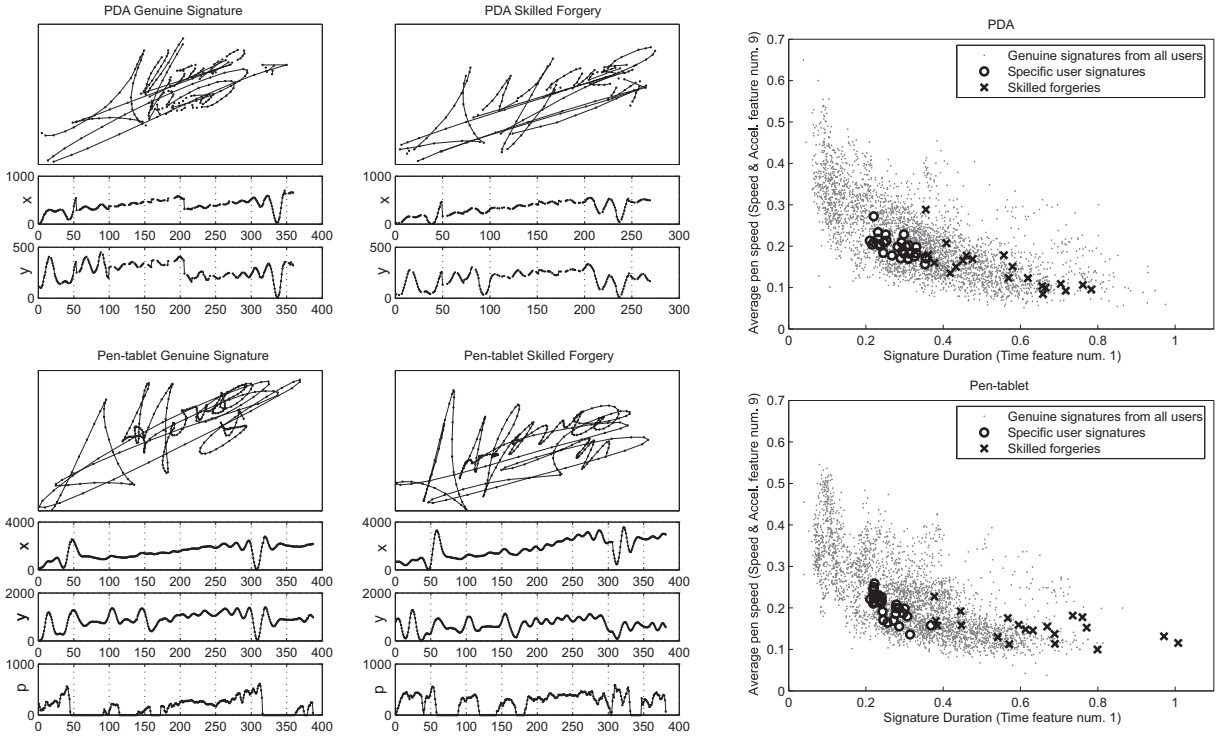


Figure 4.1: *Left: signatures from a user of the database on both scenarios and their corresponding signals used for the experiments. Azimuth and altitude signals are also available for pen tablet. No pressure signals are available for the signatures captured with PDA. Right: distribution of two example features for a particular user.*

the PDA have missing samples due to capture errors. A preprocessing step is performed in the PDA signature subcorpus to interpolate missing samples due to sampling errors. Interpolation is performed using splines (Martinez-Diaz *et al.*, 2007a), since in preliminary experimental results this leads to a better verification performance than linear interpolation. Moreover, as no pen-up information is recorded by the PDA, pen-ups are heuristically assigned wherever a gap of 50 or more milliseconds between samples exist. The samples between the pen-up and pen-down are also interpolated. A preprocessing step consisting on position normalization is performed, by aligning the center of mass of each signature to a common coordinate.

From each subset (PDA and pen tablet), the signatures of 50 users are used for development purposes, while the remaining 70 is left as a validation set to compute the verification performance. More specifically, the subset of the first 50 users is used to find the best performing features (via SFFS feature selection) and to find the optimal system parameters. The remaining 70 users are left to validate the results. This setup follows the protocol of the BioSecure Multimodal Evaluation Campaign (BMEC), where a subset of 50 users was previously released for algorithm tuning before submission to the competition (GET-INT, 2007). The evaluation was later performed by the organizers on sequestered test data corresponding to a larger set of subjects.

4.2 Experimental Protocol

User models are trained with the 5 genuine signatures from the first session, while the remaining 15 genuine signatures from the second session are left for testing.

Random forgery scores (the case where a forger uses his own signature claiming to be another user of the system) are obtained by comparing the user model to one signature sample of all the remaining users. Skilled forgery scores are computed by comparing all of the 20 available skilled forgeries per user with its own model.

4.2.1 Individual Feature Analysis

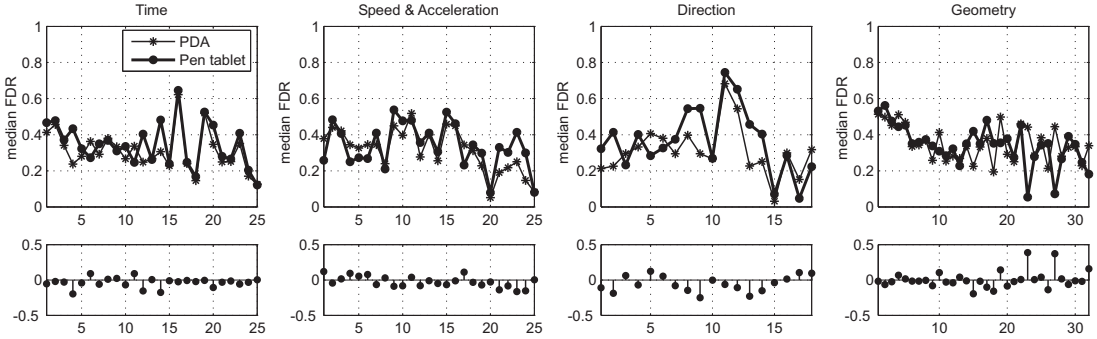
The individual discriminative power of the local and global features that are described in Chapter 3 is studied on both scenarios (PDA and pen tablet). The Fisher Discriminant Ratio (FDR) represents a suitable system-independent measure for this purpose. As a matter of fact, the FDR is a rather intuitive measure of discriminative power, as it increases with the inter-class variability and decreases with the intra-class variability. The FDR for the i-th feature of user u is computed as follows:

$$FDR_i(u) = \frac{(\mu_{G_i} - \mu_{F_i})^2}{\sigma^2_{G_i} + \sigma^2_{F_i}}$$

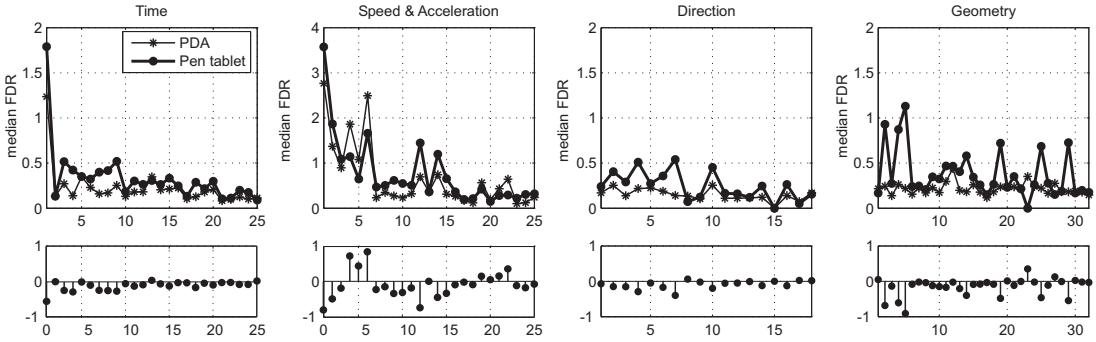
where G_i is the set of genuine signatures and F_i represents the set of forgeries. The median FDR (as the mean value is affected by outliers) over the different users in the development user set is computed to analyze the discriminative power (in terms of class separability) of each feature. The median FDR is computed differently for random and skilled forgeries. In the case of random forgeries, for each user, the FDR between the user samples and the rest of the genuine signatures in the database is computed while, for skilled forgeries, the FDR is computed between the genuine signatures of the user and the available skilled forgeries. The FDR has been also used by Lee *et al.* (1996) and Lei and Govindaraju (2005) as a measure of feature discriminative power. [1]

The Fisher Discriminant Ratio cannot be directly applied to local features, which are composed by sequences of discrete values (i.e. time functions). As a consequence, another measure must be found that preferably satisfies the system-independent property and considers the intra- and inter-class variability, which is a difficult problem. One of the main difficulties is to find a distance or dissimilarity measure between time-functions. Nalwa (1997) proposes the cross correlation between corresponding signals from two signatures as a similarity measure. However, the cross correlation is limited to sequences with the same length, so it's application to signatures is not straightforward. At this point, Dynamic Time Warping (DTW) arises as a simple and appropriate elastic distance measure between sequences with different lengths and non-linear deformations between realizations. By using DTW, we can define the following ratio for the i-th feature of user u:

[1] More specifically, the square root of the FDR is used in those works.



(a)



(b)

Figure 4.2: *Median FDR values and differences between PDA and pen tablet for random (a) and skilled forgeries (b).*

$$DDR_i(u) = \frac{(\mu_{DG_i} - \mu_{DF_i})^2}{\sigma^2_{DG_i} + \sigma^2_{DF_i}}$$

where DG_i is the set of distances between the i-th feature of the genuine signatures and DF_i is formed by the set of distances between forgeries and the genuine signatures. We call this measure the Distance Discriminative Ratio (DDR) due to its relation to the FDR. The DDR is a modified version of the consistency measure for local features proposed by Lei and Govindaraju (2005).

The EER of the global and local verification systems using each individual feature is also computed. The correlation between the individual EER and the corresponding FDR or DDR for each feature is also studied. This allows to verify the relationship between the proposed distance ratios and the system performance.

4.2.2 Feature Combination Analysis

The analysis of individual features allows to predict which types of features are likely to be part of an optimal multidimensional feature vector. Nevertheless, the existing relationships or correlation between them may alter this intuitive reasoning, and features that perform well individually may not do so in combination with others. Therefore, we perform feature selection

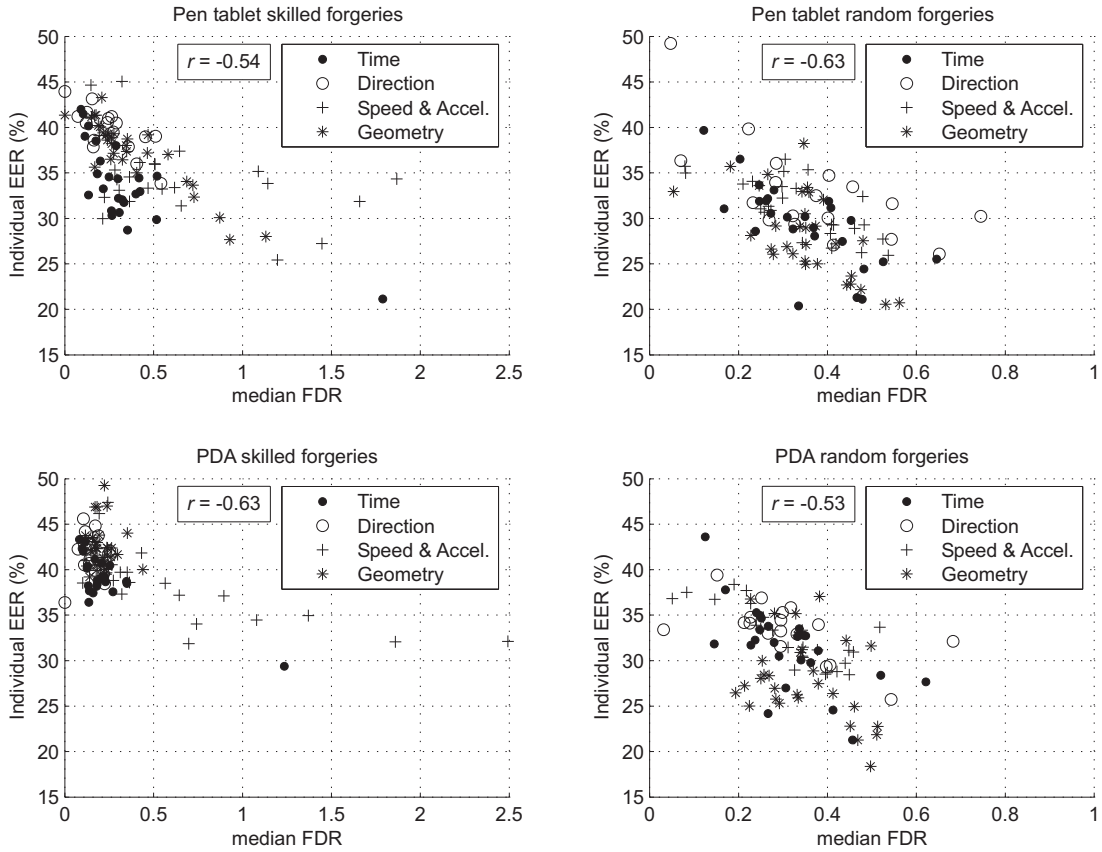


Figure 4.3: *Correlation between the median FDR and the individual EER for each feature for pen tablet (top) and PDA (bottom).*

over the whole set of proposed features to obtain optimal feature subsets for the handheld and tablet scenario. Feature selection based on the SFFS algorithm is applied separately for random and skilled forgeries. The contribution of each type of global and local features in the optimum feature set for each scenario is then studied.

Finally, the verification performance of the optimum feature subset over the validation set (the remaining 70 users) is studied, and fusion between both systems (local and global) is performed.

4.3 Development Experimental Results

4.3.1 Global Features

4.3.1.1 Individual Feature Analysis

The median FDR over all the users in the database is depicted in Fig. 4.2 for the four global feature types, as described in Sect. 3.1.

From Fig. 4.2, we observe that the median FDR for each feature is similar in the pen tablet and the PDA scenario when random forgeries are considered (top row). The individual

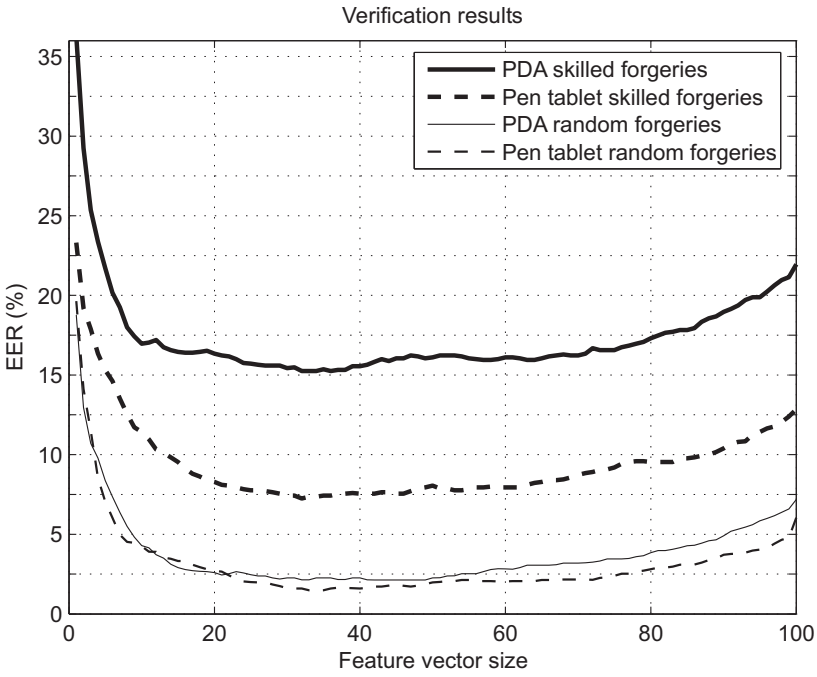


Figure 4.4: *Verification performance in terms of the size of the optimal feature selected by the SFFS algorithm.*

differences are depicted under each graph for clarification. Negative values indicate that the FDR is higher for pen tablet than for PDA, while positive values indicate the opposite. As can be seen, in most cases the FDR is higher for pen tablet, although their difference is usually small. On the contrary, it is higher for pen tablet than PDA in the case of skilled forgeries (bottom row).

These results suggest that the verification performance in the PDA scenario against skilled forgeries would be *a priori* lower than for pen tablet independently from the classifier used. The performance should be similar in the case of random forgeries.

The individual performance in terms of the EER is now computed for each feature considering random and skilled forgeries separately. The correlation between the individual EER and the median FDR can be observed in Fig. 4.3. As can be seen, the relationship between both magnitudes is clear (with relatively high values for the correlation coefficient r), proving that the median FDR is a good estimator for the discriminative power of each feature.

4.3.1.2 Feature Combination Analysis

Feature selection is performed on the development set of 50 users. In Fig. 4.4 the evolution of the system EER according to the size of the optimum feature vector selected by the SFFS algorithm is depicted. It can be seen that while the behavior for the case of random forgeries is similar on both scenarios, the verification performance is significantly better for skilled forgeries in the pen tablet scenario. If the evolution of the EER is carefully observed, it can be noticed that both plots for the PDA scenario decrease more steeply until a stable region is reached than

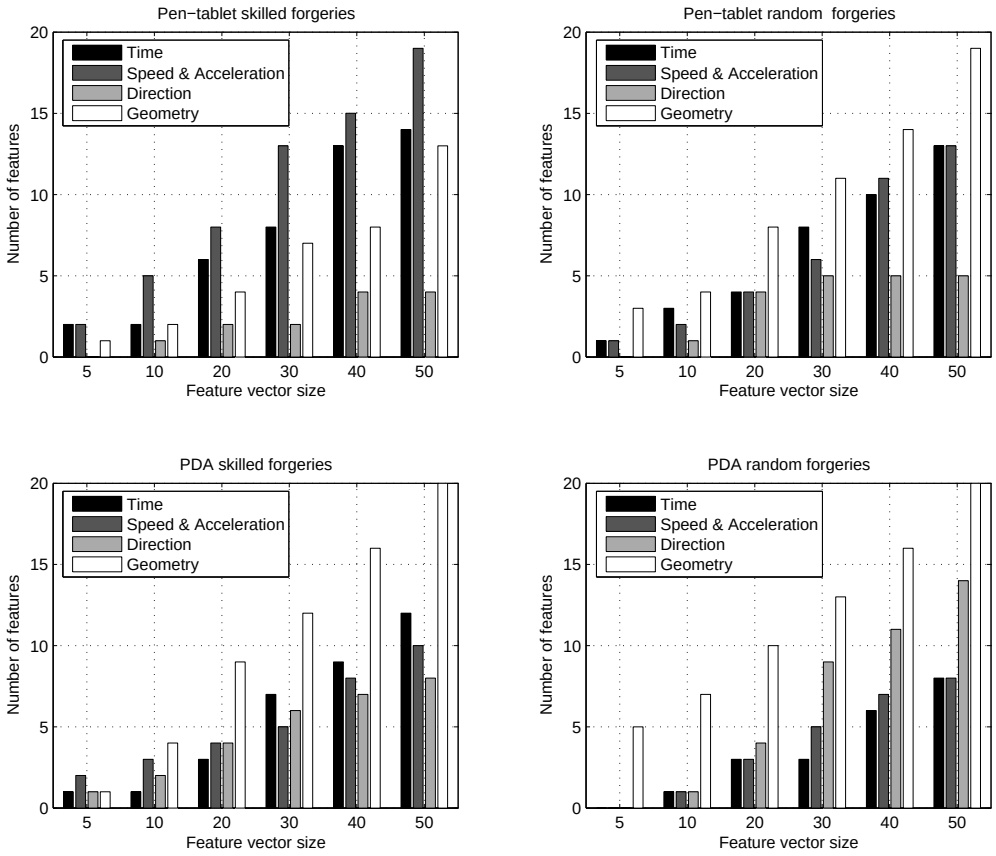


Figure 4.5: *Histogram for each type of feature for skilled (left) and random forgeries (right) on pen tablet (top) and PDA (bottom)*

their equivalent plots in the pen tablet scenario. This indicates that the near optimal feature vector for PDA requires less features than for pen tablet. Moreover, as the feature vector size increases, the performance in the PDA case degrades more rapidly than in pen tablet. These two effects suggest that PDA signatures may carry less information, that is, they can be modeled with less features and are more affected by the curse of dimensionality when more features are added.

The contribution of each type of feature in the optimal feature vector is analyzed in Fig. 4.5. An histogram of each type of feature for different sizes of the optimal feature vector computed by the SFFS algorithm is depicted for random and skilled forgeries on both scenarios. As can be seen in Fig. 4.5, Geometry features represent a very high proportion in the PDA scenario, with a reduced contribution of the rest of features. On the contrary, in the pen tablet scenario, the contribution of Geometry features is balanced with the one of Time and Speed & Acceleration features. Few Direction features are present in the optimal feature vectors for random forgeries and pen tablet. This may be due to the fact that no rotation normalization is performed in our

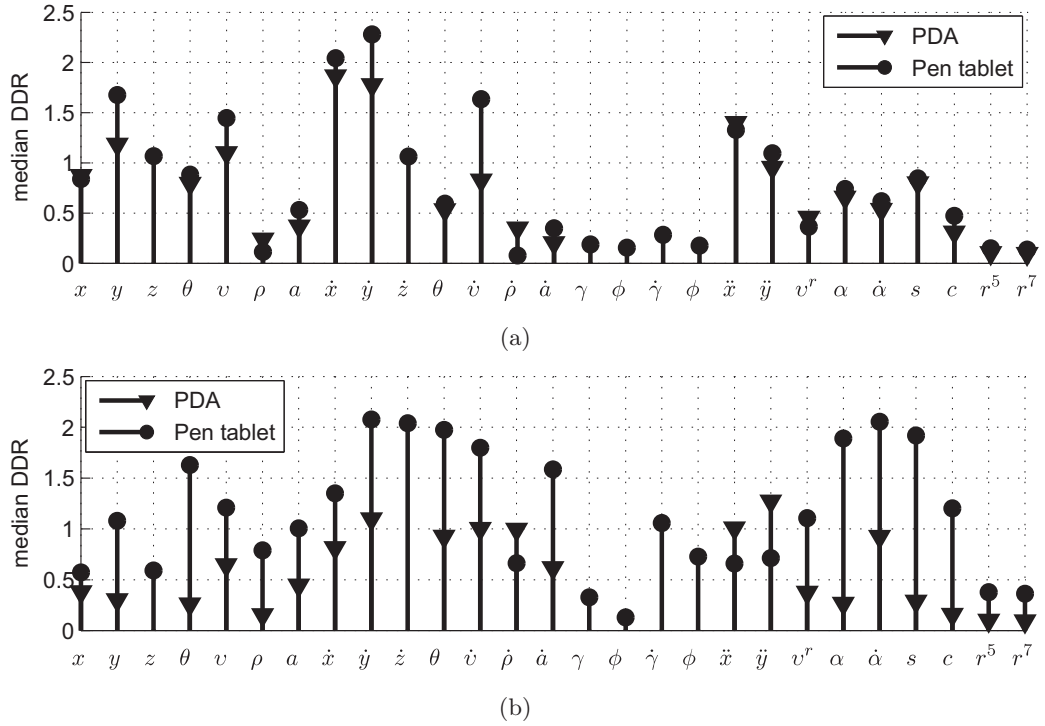


Figure 4.6: *Median DDR for the extended local feature set considering (a) random and (b) skilled forgeries. Note that some features are missing for the PDA database, since no pressure or pen inclination information is captured.*

experiments when computing the global features. Rotation may be more controlled in the PDA, as the user usually holds the PDA in a similar way. On the contrary, users may not always sign in the pen tablet with a similar orientation as their position may be more variable, depending on the particular desktop configuration found during acquisition. Speed and Acceleration features in the PDA may perform worse than for pen tablet due to the effect of sampling errors. Artificial samples are introduced by interpolation to compensate these errors, which may be affecting the discriminative power of these features.

4.3.2 Local Features

4.3.2.1 Individual Feature Analysis

To analyze local features, we follow a similar approach to the one for global features. First, the individual discriminative power based on statistical information is studied by using the Distance Discriminative Ratio (DDR) defined in Sect. 4.2.1. The DDR is computed separately for each feature, and differently for random and skilled forgeries. As in the case of global features, for random forgeries, the DDR is computed using the 5 genuine signatures of each user and one genuine signature from each of the remaining users. Similarly, for skilled forgeries, the DDR is computed using the 5 available training signatures and the 20 available corresponding forgeries. The median values for each feature described in Table 3.1 are depicted in Fig. 4.6. It must be

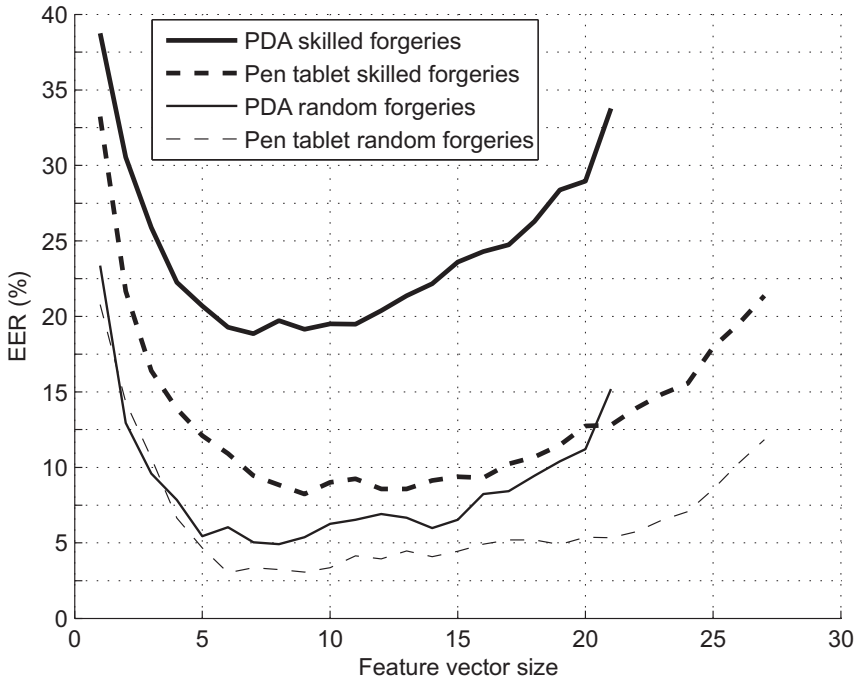


Figure 4.7: *Local system verification performance in terms of the size of the optimal feature selected by the SFFS algorithm.*

noted that some features are not available in the signatures captured using a PDA, since neither pressure nor orientation information is available in that scenario.

As can be seen, the median DDR is in general similar on pen tablet and PDA for random forgeries, while it is higher for pen tablet when skilled forgeries are considered. The obtained median DDR values are aligned with other related works. For example, Lei and Govindaraju (2005) concluded that the coordinate sequences (x, y), speed ($\dot{x}$, $\dot{y}$, v) and angle (α, $\cos(\alpha)$, $\sin(\alpha)$) have a high consistency.

4.3.2.2 Feature Combination Analysis

We take an equivalent approach than the one for global features to analyze feature selection with local features. However, the configuration of the HMM system must be carefully chosen, as a too complex HMM will heavily penalize low-dimensional feature vectors and a too simple one may lead to non representative results. We opt for the original configuration described in (Fierrez *et al.*, 2007), with 2 states and 32 Gaussian Mixtures per state.

Feature selection using the SFFS algorithm is performed over the 21 (PDA) and 27 (pen tablet) feature sets presented in Table 3.2 separately for random and skilled forgeries. As for global features, feature selection is performed on the development set of 50 users. The verification performance in terms of the EER for all the possible values of the optimal feature vector dimensionality is depicted in Fig. 4.7. As can be seen, a better performance is in general achieved in the pen tablet scenario.

Table 4.1: *Selected optimal feature vectors by the SFFS algorithm on the development set.*

Scenario	Best performing features
PDA skilled forgeries	x, y, v, ρ, $\dot{y}$, $\dot{\rho}$, c
PDA random forgeries	x, y, ρ, $\dot{x}$, $\dot{y}$, $\dot{\theta}$, $\dot{\alpha}$, c
Pen tablet skilled forgeries	x, y, v, $\dot{y}$, $\dot{\theta}$, $\dot{v}$, v^r, $\dot{\alpha}$, c
Pen tablet random forgeries	x, v, $\dot{y}$, $\dot{\theta}$, $\dot{\alpha}$, c

Table 4.2: *System performance on the validation set using global or local features on both scenarios. Vectors of 40 features have been selected in every configuration for the global system.*

Optimization scenario	**Global**		**Local**	
	$EER_{rd}(\%)$	$EER_{sk}(\%)$	$EER_{rd}(\%)$	$EER_{sk}(\%)$
PDA, skilled forgeries	7.23	16.29	6.03	17.48
PDA, random forgeries	5.43	17.72	5.79	22.24
Pen tablet, skilled forgeries	5.61	11.31	5.42	11.10
Pen tablet, random forgeries	6.69	12.98	5.74	13.80

As with global features, it can be also noted that the verification performance increases and degrades more rapidly in the PDA scenario as more features are added. It could be hypothesized that this is due to the fact that less features are considered for PDA, as neither pressure nor pen orientation information is available. However, it was found in the experiments that none of the features related to these magnitudes were selected by the SFFS algorithm in the best performing feature vector sizes. So, at least, it can be stated that less features are required in the PDA scenario to achieve the optimum performance. This corroborates the hypothesis introduced in Sect. 4.3.1, stating that signatures captured with the PDA carry less information.

The optimal feature combinations selected by the SFFS algorithm are presented in Table 4.1. Several remarks can be extracted from these results. First, neither pressure nor pen orientation-related features are present in the optimal feature vectors, suggesting that the lack of them should not penalize the verification performance. It can also be seen that no second derivative-related features are present in the optimal sets. Three features are present in all vectors, namely the x coordinate, the first derivative of the y coordinate and the cosine c of the trajectory angle α. Only 13 of the 27 proposed features are present in any of the optimal feature vectors.

These results reveal that less features are needed for HMM-based signature verification compared to the ones commonly considered in other works such as (Fierrez *et al.*, 2007; Richiardi *et al.*, 2005; Van *et al.*, 2007), at least under these experimental conditions. Contrary to the results presented by Muramatsu and Matsumoto (2007), pressure or pen-orientation features prove to be less discriminative than the rest of features (at least when they are combined with

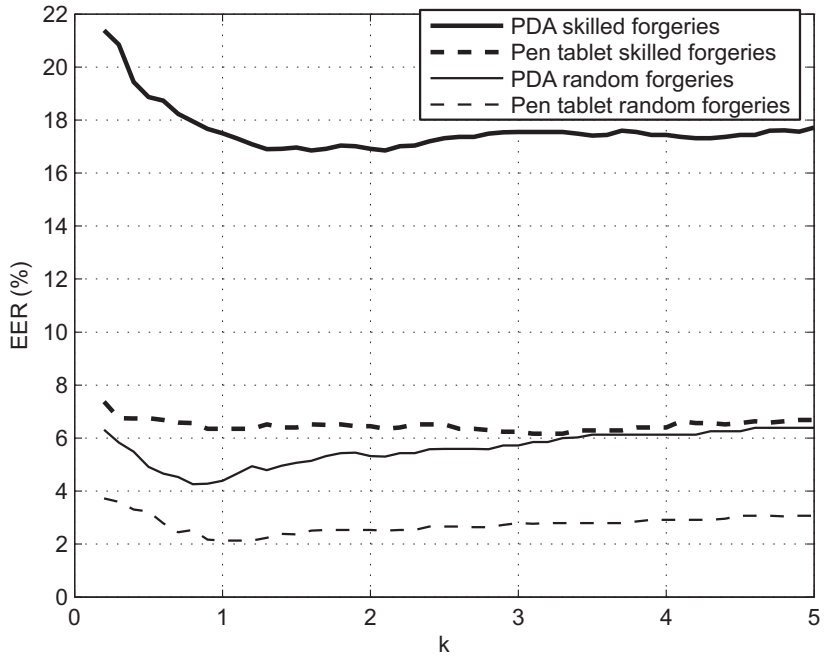
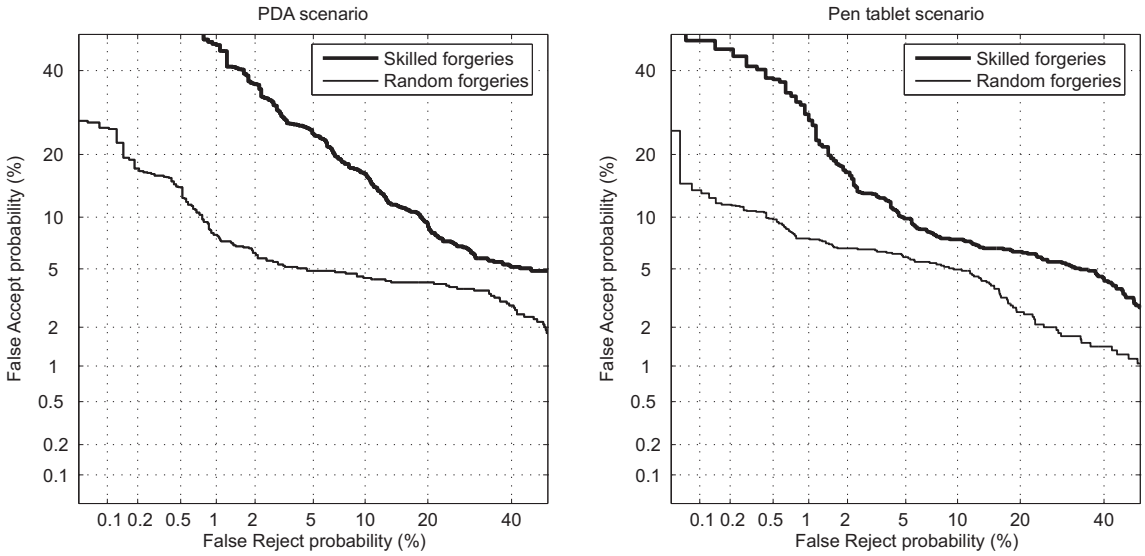


Figure 4.8: *System verification performance on the development set using score fusion for different values of the fusion weighting coefficient k.*

other types of features). The absence of pressure in the optimal feature vectors suggests that a pen tablet-based system does not have *a priori* advantage over a PDA system due to the capture of pressure information. It is also worth noting that most the best performing features selected by the algorithm are simple in terms of their computation, like the raw coordinate sequences, speed and angle-related features. Pen inclination-related features, which are not available in the PDA, are not part of the optimal feature vectors for pen tablet. In our experiments, it is observed that some of them are present in near-optimal vectors (i.e. optimal feature vectors with a dimensionality that is close to the optimal one), especially the ones related to pen-azimuth.

4.4 Validation Experimental Results

Until now, the verification performance of the selected feature subsets has been computed using the development set of 50 users. These results may be consequently affected by overfitting to the development data. To validate our results, we compute the verification performance on the test set of 70 users on the database selecting the best performing feature combinations on the development set for global and local features. Results are given in Table 4.2.



(a) (b)

Figure 4.9: *DET plots for the (a) PDA and (b) Pen tablet scenario using score fusion and feature vectors optimized against skilled forgeries.*

Table 4.3: *System performance on the validation set of 70 users using score fusion.*

Optimization Scenario	$EER_{rd}(\%)$	$EER_{sk}(\%)$
PDA, skilled forgeries	4.70	12.29
Pen tablet, skilled forgeries	5.65	7.80

4.5 Fusion of Global and Local Systems

Fusion Development Experiments

Fusion of the global and local systems with optimal feature vectors is performed via weighted sum of the match scores (Ross *et al.*, 2006). Only the systems optimized for skilled forgeries on both scenarios are considered, as they provide a reasonable performance against random forgeries. The fusion score s_f is obtained as

$$s_f = s_g + k \cdot s_l$$

where s_g and s_l are the match scores of the global and local systems respectively. The fusion weighting coefficient k is heuristically set by observing the system performance in terms of the EER on the development signature set. The system EER for different values of k on both scenarios is shown in Fig. 4.8. As can be seen, the results are considerably better for skilled forgeries in the pen tablet scenario than for PDA, primarily due to the better performance of the global and local systems against skilled forgeries using pen tablet.

An optimal value of $k = 1$ is chosen for the pen tablet scenario. The most appropriate

value of k is less clearer for the PDA scenario, as it is notably different for random and skilled forgeries. As our aim is to optimize against skilled forgeries, we set a value of $k = 1.5$ for PDA. The higher optimal value for k in the PDA scenario and skilled forgeries indicates that the local system is in this case more discriminative than the global one. On the contrary, an optimal value of k below 1 for random forgeries indicates that the global system performs better than the local one in the PDA scenario.

Fusion Validation Results

The system performance using score fusion on the validation set of 70 users is finally computed by using the heuristically selected weighing coefficient k on each scenario. The system performance is represented by DET plots in Fig. 4.9. The EER is also shown in Table 4.3. As can be observed, the verification performance is similar on both scenarios when random forgeries are considered. On the contrary, the performance is much better for pen tablet in the case of skilled forgeries, even though both systems have been tuned specifically against skilled forgeries via feature selection.

A notable increase in the performance can be observed when score fusion is applied with respect to the individual local and global systems, especially for the case of skilled forgeries.

4.6 Discussion: Pen-up trajectories

It is observed from the results that in general the discriminative power of dynamic features against skilled forgeries is lower in the PDA scenario compared to the common pen tablet acquisition. On the other hand, the results are similar in both scenarios when random forgeries are considered. The individual feature class separability measures indicate a slightly higher discriminative power for pen tablet and random forgeries, although the final verification performance is even better for PDA when considering random forgeries in some cases.

As has been previously stated, the observed differences in the performance and class separability measures may be due to the increased variability and sampling errors in the PDA scenario. Another key issue that must be taken into account is the inability of the PDA to capture signals during pen-ups. In the pen-tablet signature database used for the experiments, it is obtained that among genuine signatures an average of 18% of the captured signature signals correspond to pen-ups. The histogram of the proportion of number of pen-up samples compared to the total signature samples is presented in Fig. 4.10.

The lack of pen movement information during pen-ups may negatively affect the verification performance. In order to test this hypothesis, the signatures from the pen-tablet database are modified by eliminating the samples corresponding to pen-ups (i.e. with zero pressure). Pen-ups are interpolated using splines, following the same procedure than with the PDA dataset signatures. A new database of pen-tablet signatures from the original 120 users with interpolated pen-ups is obtained. This dataset is divided in a development set of 50 users and a validation set of 70 users, as in the previous experiments.

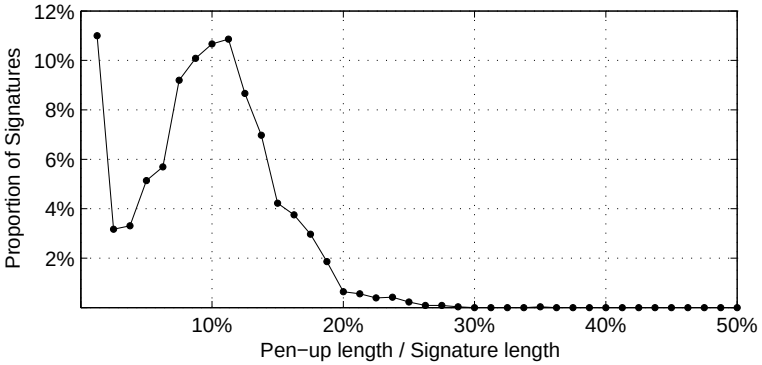


Figure 4.10: *Histogram of the proportion of pen-up samples compared to the total number of the captured signature samples in the pen tablet scenario.*

Feature selection both for global and local features is performed on the development set of 50 users. The verification performance for each possible size of the optimal feature vector for local and global features is presented in Fig. 4.11. As can be seen, the verification performance against skilled forgeries is negatively affected only in the local verification system by the lack of pen-up trajectory information. The performance is not affected in the global system. Against random forgeries, the EER remains unaffected by the lack of pen-up information.

The robustness of the global system against the lack of pen-up information may be due to the holistic nature of the extracted features. The average 18% of samples that are interpolated are not significatively affecting the feature values, or at least the relationship among the feature values from different signatures. On the other hand, the HMM-based local system is highly affected by the lack of pen-up information when skilled forgeries are considered.

These results suggest that even though pen-ups are interpolated in the PDA scenario, the verification performance is similar than in the pen tablet scenario. In the case of skilled forgeries, the error rate is much higher in the PDA scenario in part, but not only, due to the lack of pen-up trajectory information.

4.7 BioSecure Multimodal Evaluation Campaign - Mobile Scenario

The BioSecure Multimodal Evaluation Campaign (BMEC) was held in 2007 with the aim of comparing the performance of verification systems from different research groups on individual biometric modalities and fusion scenarios (Alonso-Fernandez *et al.*, 2008). In particular, the Mobile Scenario consisted on 4 modalities and fusion, using a subset of the BioSecure Multimodal Database captured on mobile conditions (i.e. using portable devices such as a PDA).

In this section we describe the systems presented to the Signature Verification modality of BMEC 2007 - Mobile Scenario and compare them to our system. In this evaluation, a signature subset from the BioSecure database was used, although it is not the same set than the one used

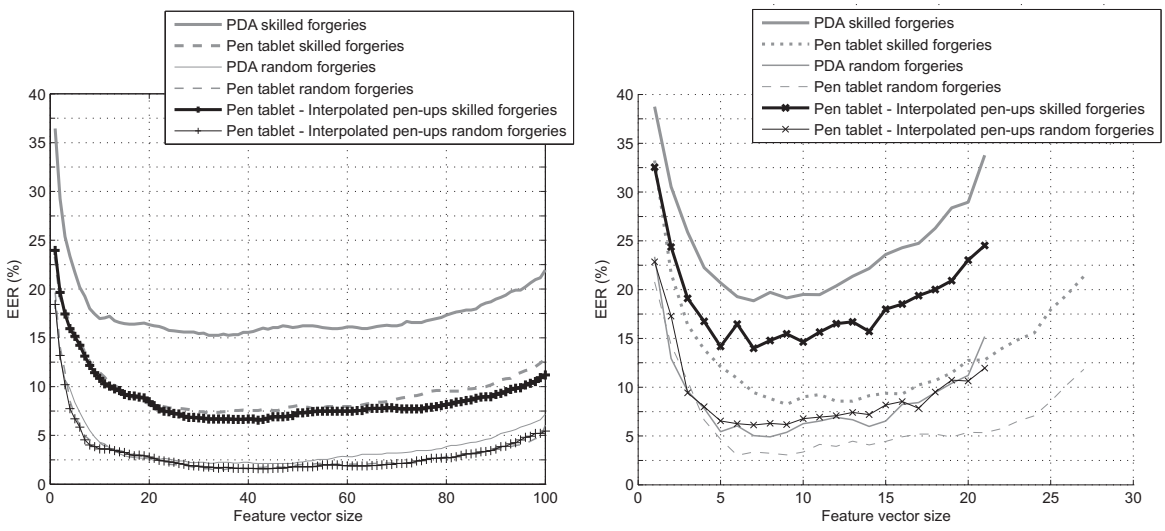


Figure 4.11: *Verification performance of the (a) global and (b) local systems in terms of the size of the optimal feature set selected by the SFFS algorithm.*

in this work. This is due to privacy issues, as the BioSecure Database has not been released yet. Nevertheless, the protocol followed in this work is equivalent, as 50 users were issued for development purposes and the evaluation was carried on an larger independent dataset (of ca. 400 subjects). As in this work, user models were computed from 5 genuine signatures.

A total of 11 systems from 6 institutions were presented. For the sake of clarity, we will refer to these systems using the same nomenclature as in the evaluation report (GET-INT, 2007). We provide a high level description of each system:

- **AMSL**: this system is based on the Levenshtein or edit distance algorithm (Schimke *et al.*, 2004). The signature is modeled as an event string, derived from the pen motion like pen-ups and the maxima and minima from position and speed. The edit distance is used to match strings extracted from the signatures.

- **EPFL1 and EPFL2**: these systems use Gaussian Mixture Models (GMM) and local features. They differ on the pre-processing steps and the number of Gaussian components (24 and 36 respectively). Scores are normalized by a world model with four Gaussian components.

- **EPFL3 and EPFL4**: these two systems are GMM multi-classifiers (7 and 6 classifiers respectively). They are composed of five GMMs using local features and two and one using global features respectively. Both systems implement a world model as EPFL1 and EPFL2 to compute the final match score. EPFL4 extracts six quality measures from the signatures. This measures are fused with the match scores to produce a binary output (accept or reject). The four EPFL systems are modified versions of those introduced in (Richiardi *et al.*, 2005).

Table 4.4: *Signature verification results of the BioSecure Multimodal Evaluation Campaign - Mobile Scenario. Results are given in terms of EER for random (rd) an skilled (sk) forgeries. The best results obtained in the present thesis following a similar experimental protocol (on a different set of signatures) are included for reference (in* ***bold****).*

System	$EER_{rd}(\%)$	$EER_{sk}(\%)$
AMSL	8.03	24.31
EPFL1	6.14	18.03
EPFL2	5.12	17.71
EPFL3	4.03	13.58
EPFL4	11.24	17.48
GET-INT	8.07	13.43
UNIFRI1	4.16	14.14
UNIFRI2	5.61	16.09
UPM1	5.18	30.39
UPM2	6.58	21.45
UniTOURS	13.14	29.14
Ref. System	4.88	15.36
Our approach: local + global. Optimized for PDA & skilled forgeries	4.70	12.29

- **Reference System and GET-INT**: a 19- and a 21-dimensional local feature vector is extracted at each point respectively to train an HMM. Information from the likelihood and the Viterbi path are fused to compute the match score, as described in (Van *et al.*, 2007). The GET-INT system performs interpolation of missing samples and user-dependent score normalization.

- **UNIFRI1**: an HMM is trained with 19 local features extracted at each sample of the signatures. A world model computed on the development dataset of 50 users is used for score normalization. This system is derived from the one presented in (Humm *et al.*, 2007).

- **UNIFRI2**: this system is based on a GMM with the same features as UNIFRI1. For each user, the GMM is adapted from a Universal Background Model (UBM) with the Maximum a Posteriori algorithm. The UBM is further employed as a world model for score normalization.

- **UPM1 and UPM2**: both are global systems relying on features selected from the 100-feature set described in Chapt. 3 by the SFFS algorithm. The Mahalanobis distance is used to compute the match score between a test signature and the user model. These systems are based on the one described in (Fierrez-Aguilar *et al.*, 2005a).

- **UniTOURS**: this system segments signatures at local speed minima. Three variations of the DTW algorithm are then applied to calculate the similarity between signatures. The final score is computed as the minimum score given by the 5 reference signatures.

As can be seen, the majority of systems rely on local features and HMMs or GMMs, although the best performing system in SVC 2004 was based on DTW (Kholmatov and Yanikoglu, 2005). The results for each system obtained in the BioSecure competition are presented in Table 4.4, together with the best results obtained in this thesis (in **bold**). It is worth noting that the global systems (UPM1 and UPM2) present a competitive performance for random forgeries, although they are outperformed by other systems in the case of skilled forgeries. This may be due to overfitting of the SFFS algorithm to the set of skilled forgeries supplied in the development dataset. The relatively high EER for all systems, especially against skilled forgeries, suggests that the Mobile Scenario signature corpus of the BioSecure Multimodal Database is a very challenging dataset. This may be primarily due to the worst case scenario approach followed for obtaining the skilled forgeries and the existing sampling errors.

Chapter 5

Conclusions and Future Work

The problem of automatic signature verification on handheld devices has been analyzed in this thesis. Signature verification as a biometric trait has been studied, and its applications on handheld devices and challenges have been outlined. The state of the art of dynamic signature verification has been summarized, focusing on feature- and function-based systems. The most recent and popular public signature databases have been described, focusing on the BioSecure Multimodal Database, which includes signatures captured on a PDA.

An analysis of the specificities of signature verification on handheld devices vs. the traditional pen tablet scenario has been performed. The comparison has been done using the BioSecure Multimodal Database which contains signatures from the same set of individuals captured with a PDA and a pen tablet, allowing a fair comparison between both scenarios. A global and a local system are considered. The global system considers the 100-feature set proposed in (Fierrez-Aguilar *et al.*, 2005a). The local system is based on the one described in (Fierrez *et al.*, 2007), to which the proposed feature set is extended with other features from the literature.

The analysis is carried out following two perspectives. First, a statistical point of view, by using statistical class separability measures of individual features. Second, an applied perspective, by optimizing the system EER via feature selection. The impact of interpolating the pen-up trajectory samples in the pen tablet scenario (imitating the PDA scenario) has also been evaluated. As a result, it is shown that for the global system, the verification performance is not affected when the pen-up trajectory is interpolated. On the other hand, the performance of the local system is negatively affected, especially against skilled forgeries.

Finally, fusion of the optimal systems against skilled forgeries (in terms of the feature selection algorithm) is performed, and the verification performance is compared to the one obtained by other systems in the recent BioSecure Multimodal Evaluation Campaign - Mobile Scenario (BMEC), with promising results. In this evaluation, several signature verification systems were compared using an equivalent protocol compared to the one used in this thesis.

5.1 Conclusions

The individual feature discriminative power as well as the verification performance of the systems considered is similar against random forgeries on both scenarios (PDA and pen tablet). In the PDA scenario, it is observed that when using feature selection, geometrical features are the majority in the optimal feature vector. This may be due to the sampling errors, which may affect the computation of features related to speed or orientation. The inability to capture the pen trajectory during pen-ups (which must be interpolated) by handheld devices is also affecting the verification performance, as less user-specific information is contained on the signatures. When performing feature selection, the better relative performance of low-dimensional optimal feature vectors compared to higher dimensional ones in the pen-tablet scenario suggests that signatures captured with the PDA carry less information compared to signatures captured with a pen tablet.

The discriminative power of the features against skilled forgeries is much higher on the pen tablet scenario. It has been observed that, for the local system, this difference is partly due to the lack of the pen trajectory information during pen-ups on the PDA, which represents in average the 18% of the total signature signal length. Another key factor in the performance variation may be the acquisition protocol for skilled forgeries on both scenarios. In the PDA scenario, forgers had on-screen access to the original strokes. This may have led to a better quality of forgeries in the PDA signature subcorpus. This hypothesis is corroborated by the similar verification performance against random forgeries on both scenarios for both systems. Moreover, the verification performance of the fusion system against random forgeries is even better for PDA than for pen tablet.

The applied feature selection algorithm has shown that less features than the traditionally considered when using HMMs for signature verification are necessary to achieve an optimal verification performance. It is worth noting the promising performance achieved by the global system, contrary to the intuitive idea that local systems are usually more reliable. This had already been observed in the results of the BioSecure Multimodal Evaluation Campaign (BMEC), where the results against random forgeries of the proposed global system were ranked among the best performing systems.

Fusion of the local and global systems improves the system EER in ca. 25% both for random and skilled forgeries. Score normalization techniques, like the ones proposed by the participants in BMEC may further improve the verification performance.

5.2 Future Work

Considering the results obtained in this work, a number of potential research lines arise. Among these, we cite the following:

- The analysis of signature verification on handheld devices using a classifier based on Dy-

namic Time Warping. This type of classifiers has reached a notable performance for signature verification (Kholmatov and Yanikoglu, 2005) with pen tablet signature databases. In most implementations in the literature, no pressure information has been used in DTW-based systems.

- The usage of quality measures (Richiardi *et al.*, 2007) to analyze the impact of the signature acquisition on mobile conditions. These measures may also be useful to perform quality-based fusion (Fierrez-Aguilar *et al.*, 2005c).

- The study of sensor interoperability (Alonso-Fernandez *et al.*, 2005) on signature verification. The signatures from the BioSecure Multimodal Database can be used to assess the feasibility of sensor interoparability between the handheld and pen tablet scenarios. This is also known as channel mismatch, which is a topic of big interest in the speaker recognition community (Ramos *et al.*, 2008).

- The proliferation of touch- and multitouch-screens in handheld devices is becoming focused in applications where no PDA stylus is required and users interface with applications using their fingertips. In this context, new possibilities such as graphical passwords (Jermyn *et al.*, 1999) or even signatures traced with the fingertips can be explored.

- Template selection and update techniques have proven to be effective for other biometric traits, such as fingerprints (Uludag *et al.*, 2004), but have not been yet studied for signature verification. Such techniques may alleviate the problem of short- and long-term variability in signature verification and the scarcity of training samples, improving the system performance in the challenging scenario of handheld devices.

- Signatures are sensitive data due to their legal implications. Consequently, user templates and captured signature data must be secured. Several techniques have been proposed in the literature to securely store biometric data (Jain *et al.*, 2008), among which some refer specifically to signatures Freire *et al.* (2008).

- The intrinsic information from signatures, and the differences in the information that is carried in the signatures captured in the scenarios considered in this thesis (PDA and pen tablet) can be studied from an Information Theory perspective. This can help to predict the relative verification performance that can be expected among the different scenarios. Some work has already been carried out on this subject (Garcia-Salicetti *et al.*).

Appendix A

Part of the work of this thesis will be presented in the 9th International Conference in Pattern Recognition, ICPR 2008. The accepted paper is appended.

Towards Mobile Authentication Using Dynamic Signature Verification: Useful Features and Performance Evaluation

Marcos Martinez-Diaz, Julian Fierrez, Javier Galbally, Javier Ortega-Garcia
Biometric Recognition Group - ATVS, EPS - Univ. Autonoma de Madrid
C/ Francisco Tomas y Valiente, 11 - Campus de Cantoblanco - 28049 Madrid, Spain
{*marcos.martinez, julian.fierrez, javier.galbally, javier.ortega*}*@uam.es*

Abstract

The proliferation of handheld devices such as PDAs and smartphones represents a new scenario for automatic signature verification. Traditionally, research on signature verification has been carried out employing signatures acquired using digitizing tablets or Tablet-PCs. In this paper we study the effects of the mobile acquisition conditions and we analyze the considerations that must be taken in the new handheld scenario. A signature verification system adapted to handheld devices via feature selection is proposed and a systematic comparison with a traditional pen tablet-based system is performed. The system is combined with another based on Hidden Markov Models using score fusion. Results confirm an increased signature variability in the case of handheld devices.[1]

1. Introduction

Despite its widespread social and legal acceptance, signature verification is still a challenging task within biometrics [1, 2]. As a behavioral biometric trait, signatures are subject to a considerable variability even on successive realizations, which can be increased over medium or large periods of time. Moreover, the possibility of creating forgeries with a relative ease, exposes a signature verification system to challenges not commonly present among other biometrics. Consequently, a signature verification system designer must face a high *intra-class* variability (between the signatures of a specific user) and a low *inter-class* variability, when forgeries are considered.

[1]This work has been supported by the Spanish Ministry of Education under project TEC2006-13141-C03-03. J. Fierrez is supported by a Marie Curie Fellowship from the European Commission. J. Galbally is supported by a FPU fellowship from the Spanish MEC.

Two main types of dynamic signature verification systems exist. *Feature-based* systems model the signature as a holistic multidimensional vector composed of global features [3]. *Function-based* systems extract time functions from the signature signal (pen coordinates, pressure, etc.) and perform signature matching via elastic or statistical techniques like Dynamic Time Warping (DTW) [4] or Hidden Markov Models (HMM) [5]. The typical architecture of an automatic signature verification system is depicted in Fig. 1.

Recently, smartphones and handheld devices have gathered a high level of popularity in the context of convergence and ubiquitous access to information and services. These devices represent a clear target for the deployment of a signature verification system, providing enough processing power and a stylus-based input. Signature verification can be used as a convenient alternative to passwords that may be forgotten or stolen for applications like e-commerce or access control. Nevertheless, signature verification on handheld devices is affected by factors not present in other input devices primarily due to a small input area, poor ergonomics or the fact that the user may be in movement. As a consequence, the signing process may be degraded.

Interestingly, the recent BioSecure Multimodal Evaluation Campaign (BMEC) [6], with the participation of independent research institutions, has shown that verification results for the case of handheld devices is significantly lower than those with other databases captured using a pen tablet [7].

In this paper, the problem of signature verification on handheld devices is studied. An analysis of the discriminant power of different types of features (temporal, geometric, etc.) is performed using the Fisher Discriminant Ratio (FDR) and feature selection algorithms. The resulting feature-based system, adapted to the handheld scenario, is further combined with an HMM system using score fusion, and the overall performance is measured against other sate-of-the-art systems using the re-

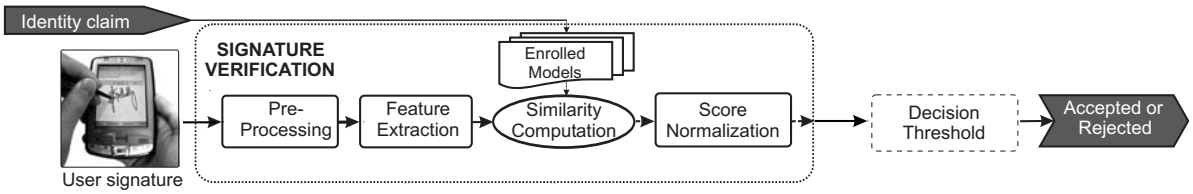


Figure 1. Signature Verification System Architecture.

sults of BMEC. The signatures used for experiments were captured both on a PDA and a digitizing tablet and correspond to the same users in both scenarios, allowing a fair comparison between them.

2. Related Work

Feature-based dynamic signature verification has been extensively studied [3, 8, 9]. Despite the large amount of different global feature sets that have been proposed (a maximum of 100 features are considered in [9]), the usually low amount of available training data motivates the use of feature selection to reduce the feature vector size (due to the curse of dimensionality). Several feature selection techniques have been proposed in the literature, being the Sequential Forward Feature Selection (SFFS) one of the best performing methods reported [10]. Nevertheless, to the extent of our knowledge, all the previous works on feature selection for dynamic signature verification have used data from digitizing tablets, so their observations may not fully apply to the case of handheld devices.

Function-based signature verification using DTW [4] or HMMs [5] is the most popular approach in signature verification. In these systems, the captured time functions are used to model each user signature. It must be taken into account that PDAs and other handheld devices are able to capture only pen position signals, while pen tablets provide additional signals such as pressure and pen inclination angles. An analysis of the implications of the lack of these signals in the PDA scenario is out of the scope of this work.

Finally, fusion of the feature- and function-based approaches has been reported to provide a better performance than the individual systems [9].

3. Signature Features

The set of features used in this work is the one presented in [9], which comprises 100 features. This is an extensive set that includes a considerable amount of features previously presented in the literature. These can be divided in four categories corresponding to the following magnitudes (the numbering is the same used in [9]):

- **Time** (25 features), related to signature duration, or timing of events such as pen-ups or local maxima: 1, 13, 22, 32, 38, 40-42, 50, 52, 58-60, 62, 64, 68, 79, 81-82, 87-90, 94, 100.
- **Speed and Acceleration** (25 features), from the first and second order time derivatives of the position time functions, like average speed or maximum speed: 4-6, 9-11, 14, 23, 26, 29, 31, 33, 39, 44-45, 48, 69, 74, 76, 80, 83, 85, 91-92, 96.
- **Direction** (18 features), extracted from the path trajectory like the starting direction or mean direction between pen-ups: 34, 51, 56-57, 61, 63, 66, 71-73, 77-78, 84, 93, 95, 97-99.
- **Geometry** (32 features), associated to the strokes or signature aspect-ratio: 2, 3, 7-8, 12, 15-21, 24-25, 27-28, 30, 35-37, 43, 46-47, 49, 53-55, 65, 67, 70, 75, 86.

Feature selection on this 100-feature set is performed using the SFFS algorithm [10], which is set to minimize the system EER using a classifier based on the Mahalanobis distance.

4. Experimental Setup

A subset of the PDA and pen tablet signature corpus of the BioSecure multimodal biometric database [6] is used for experiments. It consists of 120 users, with 20 genuine signatures and 20 skilled forgeries per user and acquisition device (PDA and pen tablet). The genuine signatures were acquired in two different sessions separated by an average period of two months, being 5 signatures from the first session and the remaining 15 from the second session. In each session, signatures were produced by the user in blocks of 5, leaving a gap of some minutes between each block. Signatures were captured with a PDA while the user was standing and holding the PDA with one hand in the handheld scenario, whereas for the pen tablet case, they were captured while the user was sitting, using a pen on a paper

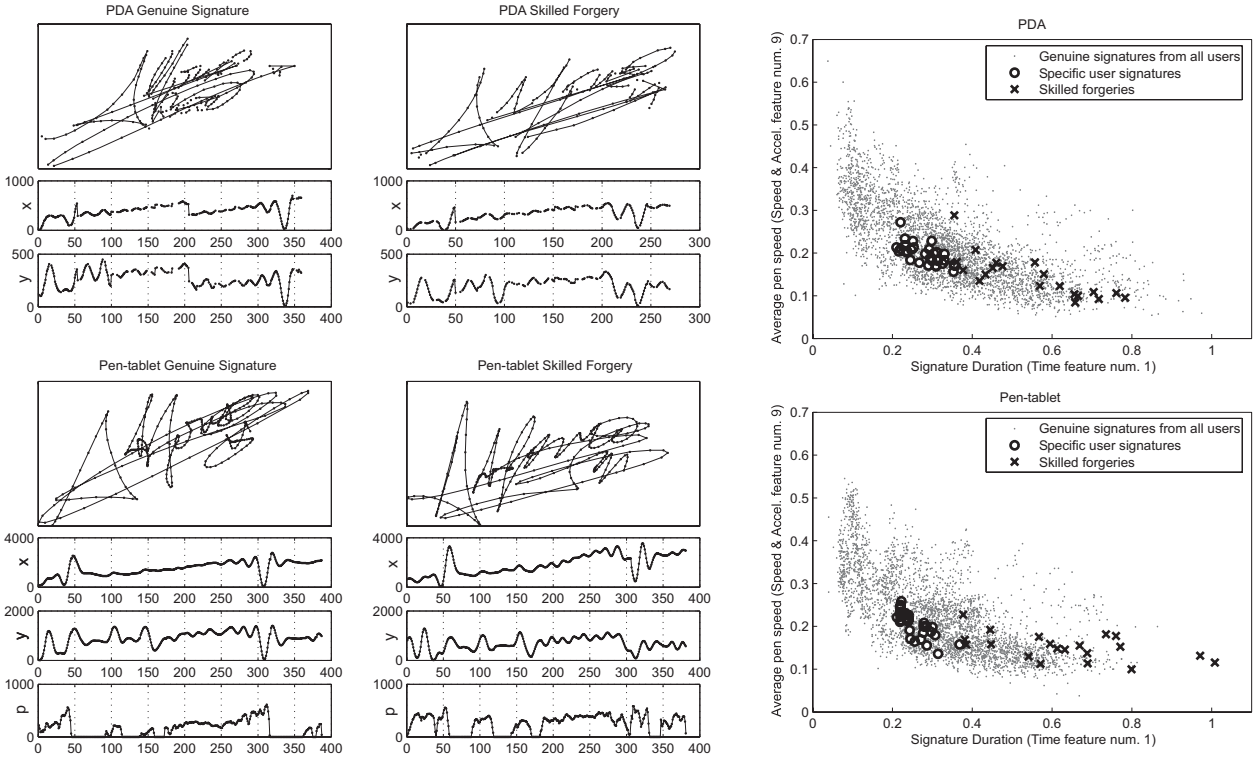


Figure 2. Left: signatures from a user of the database on both scenarios and their corresponding signals used for the experiments. No pressure signals are available for the signatures captured with PDA. Right: distribution of two example features for a particular user.

placed over the tablet. This emulates real operating conditions.

Only the x and y position signals and the sample timestamps are captured by the PDA, while pressure information and pen orientation is also provided by the pen tablet. Skilled forgeries for each user were performed by 4 different users (5 forgeries each) in a "worst case" scenario: each forger had visual access to the dynamics of the genuine signature and a tracker tool allowing to see the original strokes in both scenarios. An example of the captured signatures, their associated signals and the distribution of two features on both scenarios is shown in Fig. 2. It can be seen that signatures captured with the PDA have missing samples due to capture errors.

From each subset, the signatures from 50 users are used for development purposes, while the remaining 70 will be used to compute the verification performance. The 5 signatures from the first session are used to compute the user models. This setup follows the protocol of the BIOSECURE Multimodal Evaluation Campaign (BMEC), where a subset of 50 users was previously released for algorithm tuning before submission to the competition, which was performed by the evaluation organizers on sequestered test data. For the PDA subset, a preprocessing step is performed to interpolate missing samples.

Random forgery scores (the case where a forger uses his own signature claiming to be another user of the system) are obtained by comparing the user model to one signature sample of all the remaining users. Skilled forgery scores are computed by comparing all of the 20 available skilled forgeries per user with its own model.

The experiments are structured as follows: first, an analysis based on Fisher Discriminant Ratio (FDR) for each individual feature is performed over the development set. Next, feature selection based on the SFFS algorithm is performed (separately for random and skilled forgeries) to obtain an optimum feature subset for the handheld and tablet scenario. The contribution of each type of features (Time, Speed, etc.) in the optimum feature set for each scenario is then studied. Finally, the verification performance of the optimum feature subset over the test set (the remaining 70 users) is studied, and fusion with an HMM-based system is performed.

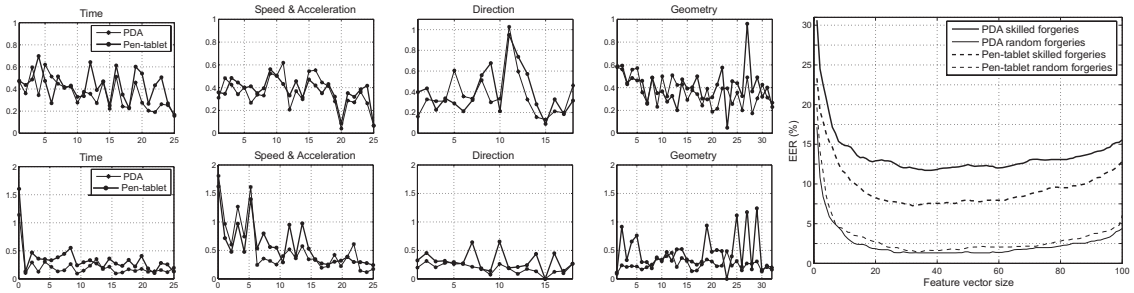


Figure 3. Median FDR over the development set for random (top) and skilled forgeries (bottom). Right: EER vs. number of features selected by the SFFS algorithm on the development set.

5. Results

To analyze the discriminative power embedded in the different feature types, the Fisher Discriminant Ratio is used as in [8], but computed for individual features and individual subjects. The median FDR (as the mean value is affected by outliers) over the different users in the development user set is depicted in Fig. 3 for the four feature types specified in Sect. 3. The FDR provides an intuitive measure of discriminative power, as it increases with the inter-class variability and decreases with the intra-class variability. The median FDR is computed differently for random and skilled forgeries. In the case of random forgeries, for each user, the FDR between the user samples and the rest of the genuine signatures in the database is computed, while for skilled forgeries, the FDR is computed between the genuine signatures of the user and the available skilled forgeries.

From Fig. 3, we observe that the median FDR for each feature is similar in the pen tablet and the PDA scenario when random forgeries are considered (top row). On the contrary, it is higher for pen tablet than PDA in the case of skilled forgeries (bottom row). This suggests that the verification performance in the PDA scenario against skilled forgeries would be *a priori* lower than for pen tablet independently from the classifier used. In Fig. 2, the distribution of the normalized values of two example features is depicted.

In Fig. 3 (right) the evolution of the system EER according to the size of the optimum feature vector selected by the SFFS algorithm is depicted. It can be observed that while the behavior for the case of random forgeries is similar on both scenarios, the verification performance is significantly better for skilled forgeries in the pen-tablet scenario.

The contribution of each type of feature is analyzed in Fig. 4. An histogram of each type of feature for different sizes of the optimal feature vector computed by the SFFS algorithm is depicted for random and skilled forgeries on both scenarios. As can be seen in Fig. 4(a) and (b), Geometry features represent a very high proportion in the PDA scenario, with a much reduced contribution of the rest of features. On the contrary, in the pen tablet scenario (Fig. 4(c) and (d)), the contribution of Geometry features is balanced with the one of Time and Speed and Acceleration features.

These results reveal that for the PDA scenario, the discriminative power of dynamic features such as Speed and Acceleration and Time features may be much lower than geometrical features. Thus, for the case of skilled forgeries, the verification performance is degraded on the PDA scenario, as Geometry features are commonly the easiest to forge.

Fusion of the PDA global feature system optimized for skilled forgeries with an user-dependent HMM system is performed via weighed sum of the match scores. The fusion weights have been heuristically adjusted to optimize both the random and skilled EERs. An optimum vector size of 50 features is selected. In the HMM system, the number of states is proportional to the mean length of the user training signatures, and the number of Gaussian Mixtures in the observations is set to maximize the likelihood of the training data with a limit of 32 mixtures. This HMM system is based on the one presented in [5]. The verification results are shown in the first row of Table 1 for the test set (70 users).

5.1. BioSecure Evaluation

The BioSecure Multimodal Evaluation Campaign [6] was held in 2007 and was composed of an Access Control Scenario and a Mobile Scenario. A signature verification modality was present in the Mobile Scenario, where signatures from the BioSecure multimodal database were used for the evaluation. A total of 11 systems were presented, from 6 independent research groups. The evaluation protocol is equivalent to the one followed in this paper (with another

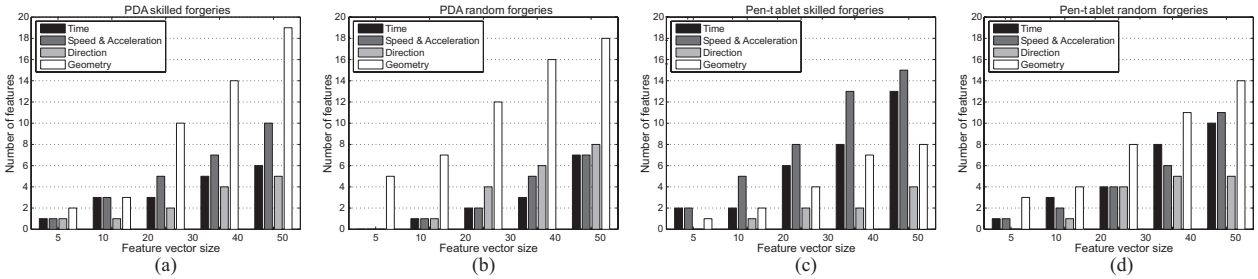


Figure 4. Histograms of feature types for different optimum feature-vector sizes and scenarios.

Table 1. PDA scenario EER comparison for random (rd) and skilled (sk) forgeries.

System	EER_{rd}	EER_{sk}
Proposed system (HMM+features)	4.0%	11.9%
BMEC best for sk. forgeries [11]	8.07%	13.43%
BMEC best for rd. forgeries [8]	4.03%	13.58%

data subset, captured in the same conditions). The best results of the evaluated systems for random and skilled forgeries are shown in Table 1, as well as the performance of the system proposed in this work. The winner system against skilled forgeries was based on an ensemble of local and global Gaussian Mixture Models and derived from [8]. The best system for random forgeries was HMM-based, using fusion of the likelihood and Viterbi path scores [11]. As can be seen, a notable verification performance has been obtained in the present work compared to those systems.

6. Conclusions and Future Work

The importance of adapting the traditional tablet-based signature verification systems to the new PDA scenario has been stated. The observed low discriminative power of dynamic features (time, speed and acceleration) in the PDA scenario suggests that ergonomics and an unfamiliar surface and signing device (touch-screen and PDA stylus vs. traditional pen and paper) may be affecting the signature process. On the other hand, the users are still able to reproduce the geometry of their own signature, which is shown by the higher consistency of geometric features. Future work includes the application of techniques aimed to compensate the increased variability found in the handheld scenario, like feature subset transformations based on session-invariant subspaces, recently introduced with significant success in the speaker recognition literature [12].

References

[1] R. Plamondon and G. Lorette. Automatic signature verification and writer identification: the state of the art. *Pattern Recognition*, 22(2):107–131, 1989.

[2] J. Fierrez and J. Ortega-Garcia. *Handbook of Biometrics*, chapter On-line signature verification. Eds. A. K. Jain, A. Ross and P. Flynn, Springer, 2008.

[3] L. L. Lee et al. Reliable on-line human signature verification systems. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(6):643–647, 1996.

[4] A. Kholmatov and B. Yanikoglu. Identity authentication using improved online signature verification method. *Pattern Recognition Letters*, 26(15):2400–2408, 2005.

[5] J. Fierrez, D. Ramos-Castro, J. Ortega-Garcia, and J. Gonzalez-Rodriguez. HMM-based on-line signature verification: feature extraction and signature modeling. *Pattern Recognition Letters*, 28(16):2325–2334, 2007.

[6] F. Alonso-Fernandez et al. Dealing with sensor interoperability in multi-biometrics: the UPM experience at the BioSecure Multimodal Evaluation 2007. In *Defense and Security Symposium, Proc. SPIE*, USA, 2008.

[7] D. Y. Yeung et al. SVC2004: First International Signature Verification Competition. In *Proc. ICBA*, pages 16–22. Springer LNCS-3072, 2004.

[8] J. Richiardi et al. Local and global feature selection for on-line signature verification. In *Proc. ICDAR*, Seoul, Korea, August-September 2005.

[9] J. Fierrez-Aguilar et al. An on-line signature verification system based on fusion of local and global information. In *Proc. AVBPA*, pages 523–532. Springer LNCS, 2005.

[10] A. K. Jain and D. Zongker. Feature selection: evaluation, application, and small sample performance. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(2):153–158, 1997.

[11] B. L. Van et al. On using the viterbi path along with HMM likelihood information for online signature verification. *IEEE Trans. on Systems, Man, and Cybernetics, Part B*, 37(5):1237 – 1247, 2007.

[12] P. Kenny et al. Speaker and session variability in GMM-based speaker verification. *IEEE Trans. on Audio, Speech and Language Processing*, 15(4):1448–1460, 2007.

Appendix B

Short Biography

Marcos Martinez-Diaz was born in 1984. He started his Telecommunication Engineering studies in 2002 and received the MSc degree in Telecommunication Engineering in 2006 from Universidad Autonoma de Madrid, Spain. He worked as an IT strategy consultant in Deloitte until 2007. Currently he is working in Business Intelligence IT project management in Vodafone.

Since 2005 he is with the Biometric Recognition Group - ATVS at the Universidad Autonoma de Madrid, where he is collaborating as a student researcher pursuing the PhD degree. His research interests include biometrics, pattern recognition and signal processing primarily focused on signature verification.

As a result of his works in the Biometric Recognition Group - ATVS, he has been awarded with the Award COIT - Bosch Security Systems to the Best MSc Thesis on Security and Surveillance Systems over Telecommunication Networks and the Honeywell Honorable Mention at the Best Student Paper Award in BTAS'07.

References

F. Alonso-Fernandez, J. Fierrez, D. Ramos, and J. Ortega-Garcia. Dealing with sensor interoperability in multi-biometrics: the UPM experience at the BioSecure Multimodal Evaluation 2007. In *Defense and Security Symposium, Biometric Technologies for Human Identification, BTHI, Proc. SPIE*, volume 6944, Orlando, USA, March 2008. 21, 29, 42

F. Alonso-Fernandez, J. Fierrez-Aguilar, F. del Valle, and J. Ortega-Garcia. On-line signature verification using Tablet PC. In *Proc. IEEE Intl. Symposium on Image and Signal Processing and Analysis, ISPA*, pages 245–250, Zagreb, Croatia, September 2005. 49

M. Ammar, Y. Yoshida, and T. Fukumura. Structural description and classification of signature images. *Pattern Recognition*, 23(7):697–710, 1990. 4

E. Anquetil and H. Bouchereau. Integration of an on-line handwriting recognition system in a smart phone device. In *Proc. of 16th Intl. Conf. on Pattern Recognition, ICPR*, volume 3, pages 192–195, 2002. 5

R. Ballagas, J. Borchers, M. Rohs, and J. Sheridan. The smart phone: a ubiquitous input device. *IEEE Pervasive Computing*, 5(1):70–77, 2006. 5

Biosecure Network of Excellence. Biosecure multimodal database. (http://www.biosecure.info), 2007. 20

CIC. http://www.cic.com. 7

Crypto-Sign. http://www.crypto-sign.com/. 7

J. Dolfing. *Handwriting Recognition and Verification, a Hidden Markov Approach.* PhD thesis, Technical University of Eindhoven, 1998. 13

J. G. A. Dolfing, E. H. L. Aarts, and J. J. G. M. van Oosterhout. On-line signature verification with Hidden Markov Models. In *Proc. of the Intl. Conf. on Pattern Recognition, ICPR*, pages 1309–1312. IEEE CS Press, 1998. vii, 10, 11, 13, 14, 18, 19

B. Dumas, C. Pugin, J. Hennebert, D. Petrovska-Delacretaz, A. Humm, F. Evequoz, R. Ingold, and D. V. Rotz. MyIDea - multimodal biometrics database, description of acquisition pro-

tocols. In *Proc. of 3rd COST 275 Workshop (COST 275)*, pages 59–62, Hatfield, UK, 2005. 22

J. Fierrez and J. Ortega-Garcia. *Advances in biometrics: sensors, systems and algorithms*, chapter Function-based on-line signature verification. Springer, 2007a. 23

J. Fierrez and J. Ortega-Garcia. *Handbook of Biometrics*, chapter On-line signature verification. Eds. A. K. Jain and A. Ross and P. Flynn, Springer, 2007b. 1, 4, 9

J. Fierrez, D. Ramos-Castro, J. Ortega-Garcia, and J. Gonzalez-Rodriguez. HMM-based on-line signature verification: feature extraction and signature modeling. *Pattern Recognition Letters*, 28(16):2325–2334, 2007. vii, 10, 11, 13, 14, 19, 20, 25, 37, 38, 47

J. Fierrez-Aguilar, N. Alonso-Hermira, G. Moreno-Marquez, and J. Ortega-Garcia. An off-line signature verification system based on fusion of local and global information. In D. Maltoni and A. K. Jain, editors, *Proc. of Intl. Workshop on Biometric Authentication, BIOAW*, pages 295–306. Springer LNCS-3087, 2004. 5

J. Fierrez-Aguilar, L. Nanni, J. Lopez-Penalba, J. Ortega-Garcia, and D. Maltoni. An on-line signature verification system based on fusion of local and global information. In *Proc. of IAPR Intl. Conf. on Audio- and Video-Based Biometric Person Authentication, AVBPA*, pages 523–532. Springer LNCS-3546, 2005a. 10, 11, 15, 16, 23, 24, 44, 47

J. Fierrez-Aguilar, J. Ortega-Garcia, and J. Gonzalez-Rodriguez. Target dependent score normalization techniques and their application to signature verification. *IEEE Trans. on Systems, Man and Cybernetics, part C*, 35(3):418–425, 2005b. 10, 25

J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, and J. Bigun. Discriminative multimodal biometric authentication based on quality measures. *Pattern Recognition*, 38(5): 777–779, 2005c. 49

M. R. Freire, J. Fierrez, and J. Ortega-Garcia. Dynamic signature verification with template protection using helper data. In *Proc. of IEEE Intl. Conf. on Acoustics, Speech and Signal Processing, ICASSP*, pages 1713–1716, March-April 2008. 49

M. Freire-Santos, J. Fierrez-Aguilar, and J. Ortega-Garcia. Cryptographic key generation using handwritten signature. In *Proc. SPIE*, volume 6202, pages 225–231, 2006. 5, 8

J. Galbally, J. Fierrez, M. R. Freire, and J. Ortega-Garcia. Feature selection based on genetic algorithms for on-line signature verification. In *Proc. of IEEE Workshop on Automatic Identification Advanced Technologies, AutoID*, pages 198–203, 2007a. 16, 17

J. Galbally, J. Fierrez, and J. Ortega-Garcia. Bayesian hill-climbing attack and its application to signature verification. In *Proc. of IAPR Intl. Conf. on Biometrics, ICB*. Springer LNCS 4642, 2007b. 8, 11

S. Garcia-Salicetti, C. Beumier, G. Chollet, B. Dorizzi, J. L.-L. Jardins, J. Lanter, Y. Ni, and D. Petrovska-Delacretaz. BIOMET: A multimodal person authentication database including face, voice, fingerprint, hand and signature modalities. In *Proc. of IAPR Intl. Conf. on Audio- and Video-based Person Authentication, AVBPA*, pages 845–853. Springer LNCS-2688, 2003. 19

S. Garcia-Salicetti, N. Houmani, and B. Dorizzi. A client-entropy measure for on-line signatures. In *Proc. of Biometrics Symposium, 2008. BSYM '08*, pages 83–88. 49

GET-INT. Biosecure multimodal evaluation campaign 2007 mobile scenario - experimental results. Technical report, 2007. (http://biometrics.it-sudparis.eu/BMEC2007/files/Results_mobile.pdf). 21, 30, 43

J. K. Guo, D. Doermann, and A. Rosenfeld. Local correspondence for detecting random forgeries. In *Proc. of the 4th Intl. Conf. on Document Analysis and Recognition, ICDAR*, pages 319–323, 1997. 4

K. Huang and H. Yan. Off-line signature verification based on geometric feature extraction and neural network classification. *Pattern Recognition*, 30(1):9–19, 1997. 5

A. Humm, J. Hennebert, and R. Ingold. Hidden Markov Models for spoken signature verification. In *Proc. of 1st IEEE Intl. Conf. on Biometrics: Theory, Applications, and Systems, BTAS 2007*, pages 1–6, 2007. 44

HYPERCOM. http://www.hypercom.com. 9

A. K. Jain, K. Nandakumar, and A. Nagar. Biometric template security. *EURASIP Journal on Advances in Signal Processing*, 2008:17 pages, 2008. Article ID 579416. 49

A. K. Jain, K. Nandakumar, and A. Ross. Score normalization in multimodal biometric systems. *Pattern Recognition*, 38(12):2270–2285, 2005. 23

A. K. Jain, A. Ross, and S. Pankanti. Biometrics: a tool for information security. *IEEE Trans. on Information Forensics and Security*, 1(2):125–143, 2006. 1

A. K. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. *IEEE Trans. on Circuits and Systems for Video Technology*, 14(1):4–20, 2004. 1, 2, 4

A. K. Jain and D. Zongker. Feature selection: evaluation, application, and small sample performance. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(2):153–158, 1997. 11, 15, 16, 18

I. Jermyn, A. Mayer, F. Monrose, M. K. Reiter, and A. D. Rubin. The design and analysis of graphical passwords. In *Proc of 8th USENIX Security Symposium*, 1999. 49

R. S. Kashi, J. Hu, W. L. Nelson, and W. Turin. On-line handwritten signature verification using Hidden Markov Model features. In *Proc. of the 4th Intl. Conf. on Document Analysis and Recognition, ICDAR*, volume 1, pages 253–257. IEEE CS Press, 1997. 14

KeCrypt. http://www.kecrypt.com. 7

R. Khokhar. Smartphones - a call for better safety on the move. *Network Security*, 2006(4):6–7, 2006. 8

A. Kholmatov and B. Yanikoglu. Identity authentication using improved online signature verification method. *Pattern Recognition Letters*, 26(15):2400–2408, 2005. 10, 11, 12, 45, 49

A. Kholmatov and B. Yanikoglu. Susig: an on-line signature database, associated protocols and benchmark results. *Pattern Analysis & Applications*, 2008. 22

F. Leclerc and R. Plamondon. Automatic signature verification: the state of the art-1989-1993. *Intl. Journal of Pattern Recognition and Artificial Intelligence*, 8(3):643–660, 1994. 4

L. L. Lee, T. Berger, and E. Aviczer. Reliable on-line human signature verification systems. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(6):643–647, 1996. 9, 10, 11, 15, 23, 31

H. Lei and V. Govindaraju. A comparative study on the consistency of features in on-line signature verification. *Pattern Recognition Letters*, 26(15):2483–2489, 2005. 10, 11, 25, 31, 32, 37

E. Maiorana, M. Martinez-Diaz, P. Campisi, J. Ortega-Garcia, and A. Neri. Template protection for HMM-based on-line signature authentication. In *(to appear) Proc. IEEE Computer Society Workshop on Biometrics, CVPR*, Alaska, 2008. 8

R. Martens and L. Claesen. Dynamic programming optimisation for on-line signature verification. In *Proc. 4th. Intl. Conf. on Document Analysis and Recognition, ICDAR*, volume 2, pages 653 – 656, 1997. 9, 10, 11, 12

A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M. Przybocki. The DET curve in assessment of decision task performance. In *Proc. of ESCA Eur. Conf. on Speech Comm. and Tech., EUROSPEECH*, pages 1895–1898, 1997. 2

M. Martinez-Diaz, J. Fierrez, M. R. Freire, and J. Ortega-Garcia. On the effects of sampling rate and interpolation in HMM-based dynamic signature verification. In *Proc. Intl. Conf. on Document Analysis and Recognition, ICDAR, IEEE Press*, volume 2, pages 1113–1117, 2007a. 30

M. Martinez-Diaz, J. Fierrez, and J. Ortega-Garcia. Universal background models for dynamic signature verification. In *Proc. IEEE Conference on Biometrics: Theory, Applications and Systems, BTAS*, pages 1–6, September 2007b. 10, 11

M. E. Munich and P. Perona. Visual identification by signature tracking. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(2):200–217, 2003. 18

D. Muramatsu and T. Matsumoto. An HMM signature verifier incorporating signature trajectories. In *Proc. of Intl. Conf. on Document Analysis and Recognition, ICDAR*, volume 1, pages 438–442. IEEE Press, 2003. 13

D. Muramatsu and T. Matsumoto. Effectiveness of pen pressure, azimuth, and altitude features for online signature verification. In *Proc. of IAPR Intl. Conf. on Biometrics, ICB*. Springer LNCS 4642, 2007. 8, 11, 38

V. S. Nalwa. Automatic on-line signature verification. *Proceedings of the IEEE*, 85(2):215–239, 1997. 31

L. Nanni and A. Lumini. Ensemble of parzen window classifiers for on-line signature verification. *Neurocomputing*, 68:217–224, 2005. 10

W. Nelson and E. Kishon. Use of dynamic features for signature verification. In *Proc. of IEEE Intl. Conf. on Systems, Man, and Cybernetics*, volume 1, pages 201–205, 1991. 23

W. Nelson, W. Turin, and T. Hastie. Statistical methods for on-line signature verification. *Intl. Journal of Pattern Recognition and Artificial Intelligence*, 8(3):749–770, 1994. 10, 23

Nemophila. http://www.nemophila.com. 7

J. Ortega-Garcia, Fierrez-Aguilar, *et al.* MCYT baseline corpus: a bimodal biometric database. *IEE Proc. Vision, Image and Signal Processing*, 150(6):391–401, 2003. 18, 21

PdaLok. http://www.pdalok.com. 7

R. Plamondon and G. Lorette. Automatic signature verification and writer identification: the state of the art. *Pattern Recognition*, 22(2):107–131, 1989. 1, 4, 5, 9

R. Plamondon and S. N. Srihari. On-line and off-line handwriting recognition: a comprehensive survey. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22:63–84, 2000. 4

P. Pudil, J. Novovicova, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15:1119–1125, 1994. 11, 16

L. R. Rabiner. A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989. 13, 14

D. Ramos, J. Gonzalez-Rodriguez, J. Gonzalez-Dominguez, and J. J. L. molina. Addressing database mismatch in forensic speaker recognition with Ahumada III: a public real-case database in Spanish. In *Proc. of Interspeech 2008*, September 2008. 49

J. Richiardi and A. Drygajlo. Gaussian Mixture Models for on-line signature verification. In *Proc. of ACM SIGMM Workshop on Biometric Methods and Applications, WBMA*, pages 115–122, 2003. 10, 13, 14

J. Richiardi, H. Ketabdar, and A. Drygajlo. Local and global feature selection for on-line signature verification. In *Proc. IAPR 8th International Conference on Document Analysis and Recognition, ICDAR*, Seoul, Korea, August-September 2005. 10, 16, 25, 38, 43

J. Richiardi, K. Kryszczuk, and A. Drygajlo. Quality measures in unimodal and multimodal biometric verification. In *Proc. of 15th European Signal Processing Conf. (EUSIPCO)*, 2007. 49

A. Ross, K. Nandakumar, and A. K. Jain. *Handbook of Multibiometrics*. Springer, 2006. 10, 40

R. Sabourin. *Off-line signature verification: recent advances and perspectives*, volume 1339, chapter Lecture Notes in Computer Science, LNCS, pages 84–98. Springer, 1997. 4

H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 26:43–49, 1978. 11

Y. Sato and K. Kogure. Online signature verification based on shape, motion and writing pressure. In *Proc. of 6th Intl. Conf. on Pattern Recognition*, pages 823–826, 1982. 10, 11, 12

S. Schimke, C. Vielhauer, and J. Dittmann. Using adapted Levenshtein distance for on-line signature authentication. In *Proc. of 17th Intl. Conf. on Pattern Recognition 2004, ICPR*, volume 2, pages 931– 934, 2004. 43

SignPlus. http://www.app-davos.ch/. 7

SOFTPRO. http://www.signplus.com. 7

S. Theodoridis and K. Koutroumbas. *Pattern Recognition*. Academic Press, 2006. 15, 16, 17

Topaz. http://www.topazsystems.com/. 7

U. Uludag, A. Ross, and A. Jain. Biometric template selection and update: a case study in fingerprints. *Pattern Recognition*, 37(7):1533–1542, 2004. 49

B. L. Van, S. Garcia-Salicetti, and B. Dorizzi. On using the Viterbi path along with HMM likelihood information for online signature verification. *IEEE Trans. on Systems, Man, and Cybernetics, Part B*, 37(5):1237 – 1247, 2007. 10, 11, 13, 14, 25, 38, 44

WACOM. http://www.wacom.com. 9

Xyzmo. http://www.xyzmo.com. 7

L. Yang, B. K. Widjaja, and R. Prasad. Application of Hidden Markov Models for signature verification. *Pattern Recognition*, 28(2):161–170, 1995. 11, 13, 14

M. Yasuhara and M. Oka. Signature verification experiment based on nonlinear time alignment: a feasibility study. *IEEE Trans. on Systems, Man and Cybernetics, part C*, 12(3):212–216, 1977. 11

D. Y. Yeung, H. Chang, Y. Xiong, S. George, R. Kashi, T. Matsumoto, and G. Rigoll. SVC2004: First international signature verification competition. In *Proc. of Intl. Conf. on Biometric Authentication, ICBA*, pages 16–22. Springer LNCS-3072, 2004. 12, 20, 25